

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2006, 01:16:00 ; Search time 194 Seconds
(without alignments)
1121.830 Million cell updates/sec

Title: US-10-679-362-2

Perfect score: 2469

Sequence: 1 MSTQRLRNEDYHYSSTDVS.....YELIQSNAPIFINSTCAFI 476

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_8:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*
- 9: geneseqp2005s:*
- 10: geneseqp2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2469	100.0	476	7	Adf90164 Human tra
2	2469	100.0	476	8	Adr03799 Human nov
3	2466	99.9	507	6	Abg75834 Transport
4	2461	99.7	476	6	Abp96441 Human tra
5	2460	99.6	496	6	Abu52588 Human NOV
6	2402.5	97.3	568	5	Aaol14191 Human tra
7	2398	97.1	500	6	Abu52587 Human NOV
8	2324	94.1	476	4	Abb76941 Human V05
9	2324	94.1	476	7	Adf90169 Human tra
10	2324	94.1	476	8	Adr03804 Human hom
11	2217	89.8	434	6	Ada55508 Human pro
12	2217	89.8	434	8	Adsl11015 Human the
13	2146.5	86.9	475	6	Abp96446 Rat LYAAT
14	2144.5	86.9	475	6	Abp96443 Mouse tra
15	2142.5	86.8	475	4	Abb76935 Rat VG51.
16	1723	69.8	382	4	Abb76934 Rat VG51.
17	1722.5	69.8	483	5	Abb98199 Human HIP
18	1722.5	69.8	483	6	Abp96437 Human tra
19	1718.5	69.6	483	5	Aae26625 Human tra
20	1718.5	69.6	507	6	Adm33124 Human tra
21	1718.5	69.6	507	6	Aae29907 Human tra
22	1713.5	69.4	483	8	Adg65947 Novel hum
23	1696.5	68.7	478	6	Abp96435 Mouse tra

24	1677.5	67.9	481	6	ABP96436	Abp96436 Rat trand
25	1544	62.5	450	5	ABB98200	Abb98200 Human HIP
26	1539	62.3	450	8	ADJ27216	Adj27216 Human TRI
27	1476.5	59.8	470	6	ABP96438	Abp96438 Human tra
28	1476.5	59.8	470	6	ADR09715	Adr09715 Human pro
29	1463	59.3	433	8	ADP29500	Adp29500 Human sec
30	1461	59.2	431	8	ADP29656	Adp29656 Human sec
31	1448.5	58.7	427	5	AAE25571	Aae25571 Human tra
32	1416.5	57.4	477	6	ABP96442	Abp96442 Mouse tra
33	1375.5	55.7	393	8	ADS11014	Adsl11014 Human the
34	1301	52.7	455	8	ADQ67776	Adq67776 Novel hum
35	1293.5	52.4	500	4	ABB76931	Abb76931 Rat VGL
36	1263	51.2	263	6	ABU52589	Abu52589 Human NOV
37	1263	51.2	322	8	ADS12237	Adsl12237 Human the
38	1263	51.2	322	8	ADS12238	Adsl12238 Human the
39	1255.5	50.9	485	6	ABB99355	Abb99355 Amino aci
40	1255.5	50.9	485	7	ADD37511	Add37511 Human tra
41	1255.5	50.9	485	10	AEF18856	Aef18856 Human Ami
42	1255.5	50.9	504	4	AAm38898	Aam38898 Human pol
43	1255.5	50.9	504	7	ADN95120	Adn95120 Human LBC
44	1255.5	50.9	516	4	AAm40684	Aam40684 Human pol
45	1252.5	50.7	476	4	AAE04910	Aae04910 Human tra

ALIGNMENTS

RESULT 1

ADP90164
ID ADF90164 standard; protein; 476 AA.
XX
AC ADF90164;
XX
DT 26-FEB-2004 (first entry)
XX
DE Human transporter protein.
XX
KW peptide therapy; human; transporter.
XX
OS Homo sapiens.
XX
PN US2003170778-A1.
XX
PD 11-SEP-2003.
XX
PF 14-MAR-2001; 2001US-00805456.
XX
PR 22-DEC-2000; 2000US-0257175P.
XX
PA (WEIM/) WEI M.
PA (YANC/) YAN C.
PA (MERK/) MERKLOV G V.
PA (KETC/) KETCHUM K A.
PA (DFRA/) DI FRANCESCO V.
XX (BEAS/) BEASLEY E M.
PI
XX Wei M, Yan C, Merklov GV, Ketchum KA, Di Francesco V, Beasley EM;
XX WPI; 2003-898259/82.
DR N-PSDB; ADF90163.
XX
PT New human transporter peptide, useful for preparing a composition for
XX treating a disease or condition mediated by human transporters.
PS Claim 1; SEQ ID NO 2; 94pp; English.

XX
CC The invention describes an isolated peptide comprising a 476 amino acid
CC sequence (S1) or its allelic variant, orthologue or fragment, where the
CC allelic variant or orthologue is encoded by a nucleic acid that
CC hybridises under stringent conditions to the opposite strand of a nucleic
CC acid comprising a sequence having 2093 or 46649 base pairs, and the
CC fragment comprises 10 contiguous amino acids of S1. The peptide is useful
CC for preparing a composition for treating a disease or condition mediated

CC by a human transporter protein. This is the amino acid sequence of a
CC human transporter protein.
XX
SQ Sequence 476 AA;

Query Match 100.0%; Score 2469; DB 7; Length 476;
Best Local Similarity 100.0%; Pred. No. 7.4e-258;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSTQRLRNEDYHSDYSDVSPSESGNLNLSPGSYQRFQSGNSTTWFTQTLIHLKGN 60
Db 1 MSTQRLRNEDYHSDYSDVSPSESGNLNLSPGSYQRFQSGNSTTWFTQTLIHLKGN 60
Qy 61 IGTGLLGLPLAVKNAGIVMGPIISLLIIGIVAVHCMGILVKCAHHFCRRLNKSFDYGDIV 120
Db 61 IGTGLLGLPLAVKNAGIVMGPIISLLIIGIVAVHCMGILVKCAHHFCRRLNKSFDYGDIV 120
Qy 121 MYGLESPCSWLNRHNAHWGRVVDFFLIIVTQLGFCVYFVFLADNFKQVIEAANGTTNNC 180
Db 121 MYGLESPCSWLNRHNAHWGRVVDFFLIIVTQLGFCVYFVFLADNFKQVIEAANGTTNNC 180
Qy 181 HNNETVILPTMDSRLYMLSFPLVLLVFIIRNLRLSIFSLLANITMLVSLVMYQFIV 240
Db 181 HNNETVILPTMDSRLYMLSFPLVLLVFIIRNLRLSIFSLLANITMLVSLVMYQFIV 240
Qy 241 QRIPDPSHLPLVAPWKTYPLFFGTAIFSFEGIGMVLPLENKMKDPKRFPLILYLGWVIT 300
Db 241 QRIPDPSHLPLVAPWKTYPLFFGTAIFSFEGIGMVLPLENKMKDPKRFPLILYLGWVIT 300
Qy 301 ILVYISGLCYLQFGANIQSITLNLPCNWLQSVKLLYSIGIFFFTYALQFYVPAEIIIP 360
Db 301 ILVYISGLCYLQFGANIQSITLNLPCNWLQSVKLLYSIGIFFFTYALQFYVPAEIIIP 360
Qy 361 FFVSRAPHECELVDLFRVTVLVCILAILPRDLVLSVSVSSSALALIIPPLE 420
Db 361 FFVSRAPHECELVDLFRVTVLVCILAILPRDLVLSVSVSSSALALIIPPLE 420
Qy 421 VTTYFSEGMSPLTIFKDALISILGFVGVVGTVEALYELIQPSNAPIFINSTCAFI 476
Db 421 VTTYFSEGMSPLTIFKDALISILGFVGVVGTVEALYELIQPSNAPIFINSTCAFI 476

RESULT 2
ADRO3799
ID ADRO3799 standard; protein; 476 AA.
XX
AC ADRO3799;
XX
XX 21-OCT-2004 (first entry)
XX Human novel transporter protein.
XX Human; chromosome 5; transporter.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Misc-difference 431..432
XX /note= "Encoded by CCCCCTC (SEQ ID 3 only)"
XX
XX US2004146887-A1.
XX
XX 29-JUL-2004.
XX
XX 07-OCT-2003; 2003US-00679362.
XX
XX 22-DEC-2000; 2000US-0257175P.
XX 14-MAR-2001; 2001US-00805456.
XX
XX (APPL-) APPLERA CORP.
XX
XX Wei M, Yan C, Merklov G, Ketchum KA, Difrancesco V, Beasley EM;
XX

DR WPI; 2004-552659/53.
DR N-PSDB; ADRO3798, ADRO3800.
XX
XX New nucleic acid encoding human transporter peptides, useful for
PT preparing agents for treating a disease or condition mediated by human
PT transporters.
XX
XX Claim 12; SEQ ID NO 2; 62pp; English.
XX
XX The invention relates to a new isolated nucleic acid comprising a
CC sequence encoding a transporter protein (appearing as ADRO3799) and
CC appearing as ADRO3798 (the cDNA) or ADRO3800 (the gene). Also included
CC are a vector comprising the isolated nucleic acid, a host cell containing
CC the vector, producing the polypeptide, detecting the presence of the
CC nucleic acid in a sample and an isolated human peptide having a sequence
CC that shares at least 70% homology with ADRO3799. The gene encoding the
CC human transporter is located on chromosome 5. The peptide is useful for
CC preparing agents for treating a disease or condition mediated by a human
CC transporter protein. The present sequence represents the transporter
CC protein.
XX
SQ Sequence 476 AA;

Query Match 100.0%; Score 2469; DB 8; Length 476;
Best Local Similarity 100.0%; Pred. No. 7.4e-258;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSTQRLRNEDYHSDYSDVSPSESGNLNLSPGSYQRFQSGNSTTWFTQTLIHLKGN 60
Db 1 MSTQRLRNEDYHSDYSDVSPSESGNLNLSPGSYQRFQSGNSTTWFTQTLIHLKGN 60
Qy 61 IGTGLLGLPLAVKNAGIVMGPIISLLIIGIVAVHCMGILVKCAHHFCRRLNKSFDYGDIV 120
Db 61 IGTGLLGLPLAVKNAGIVMGPIISLLIIGIVAVHCMGILVKCAHHFCRRLNKSFDYGDIV 120
Qy 121 MYGLESPCSWLNRHNAHWGRVVDFFLIIVTQLGFCVYFVFLADNFKQVIEAANGTTNNC 180
Db 121 MYGLESPCSWLNRHNAHWGRVVDFFLIIVTQLGFCVYFVFLADNFKQVIEAANGTTNNC 180
Qy 181 HNNETVILPTMDSRLYMLSFPLVLLVFIIRNLRLSIFSLLANITMLVSLVMYQFIV 240
Db 181 HNNETVILPTMDSRLYMLSFPLVLLVFIIRNLRLSIFSLLANITMLVSLVMYQFIV 240
Qy 241 QRIPDPSHLPLVAPWKTYPLFFGTAIFSFEGIGMVLPLENKMKDPKRFPLILYLGWVIT 300
Db 241 QRIPDPSHLPLVAPWKTYPLFFGTAIFSFEGIGMVLPLENKMKDPKRFPLILYLGWVIT 300
Qy 301 ILVYISGLCYLQFGANIQSITLNLPCNWLQSVKLLYSIGIFFFTYALQFYVPAEIIIP 360
Db 301 ILVYISGLCYLQFGANIQSITLNLPCNWLQSVKLLYSIGIFFFTYALQFYVPAEIIIP 360
Qy 361 FFVSRAPHECELVDLFRVTVLVCILAILPRDLVLSVSVSSSALALIIPPLE 420
Db 361 FFVSRAPHECELVDLFRVTVLVCILAILPRDLVLSVSVSSSALALIIPPLE 420
Qy 421 VTTYFSEGMSPLTIFKDALISILGFVGVVGTVEALYELIQPSNAPIFINSTCAFI 476
Db 421 VTTYFSEGMSPLTIFKDALISILGFVGVVGTVEALYELIQPSNAPIFINSTCAFI 476

RESULT 3
ABG75834
ID ABG75834 standard; protein; 507 AA.
XX
XX AC ABG75834;
XX
XX DT 10-MAY-2003 (first entry)
XX
XX DE Transporters and ion channels protein 16, TRICH-16.
XX
XX Human; transporter and ion channel; TRICH; Gene therapy;
XX cell proliferative disorder; transport disorder; neurological disorder;
XX muscle disorder; immunological disorder; amyotrophic lateral sclerosis;
XX

XX cysic fibrosis; diabetes; Parkinson's disease; prostate cancer;
KW cardiac disorder; angina; Alzheimer's disease; amnesia; epilepsy;
KW schizophrenia; sickle cell anaemia; infertility; hyperglycaemia;
KW hypoglycaemia; hypercholesterolaemia; stroke; multiple sclerosis;
KW motor neuron disorder; prion disease; metabolic disease;
KW developmental disorder; central nervous system; cardiomyopathy;
KW hypertension; asthma; AIDS; allergy; anaemia; atherosclerosis;
KW atopic dermatitis; diabetes mellitus; osteoarthritis; osteoporosis;
KW rheumatoid arthritis; psoriasis; infection; trauma; hepatitis; cancer;
leukemia; lymphoma.

XX Homo sapiens.
OS
XX WO2003016493-A2.
PN
XX 27-FEB-2003.
PD
XX

PF 16-AUG-2002; 2002WO-US026323.
XX
PR 17-AUG-2001; 2001US-0313242P.
XX
PR 21-SEP-2001; 2001US-0324782P.
PR
PR 02-OCT-2001; 2001US-0328184P.
PR
PR 26-OCT-2001; 2001US-0345937P.
PR
PR 01-NOV-2001; 2001US-0335698P.
PR
PR 13-NOV-2001; 2001US-0332804P.
PR
PR 27-NOV-2001; 2001US-0333922P.
PR
PR 26-APR-2002; 2002US-0375637P.
PR
PR 03-MAY-2002; 2002US-0377444P.
PR
PR 11-JUN-2002; 2002US-0388180P.
XX

PA (INCY-) INCYTE GENOMICS INC.
XX
XX Tang YT, Lal PG, Yue H, Baughn MR, Nguyen DB, Yao MG, Greene BD;
PI Borowsky ML, Lee S, Emerling BM, Xu Y, Becha SD, Gorvad AE;
PI Azimzai Y, Yue H, Elliott VS, Lee EA, Yang J, Lehr-Mason PM;
PI Runkumar J, Lee SY, Faris M, Turner C, Furness M, Buchbinder JL;
PI Wallia NK, Li JX, Forsythe IJ, Griffin JA, Gietzen KJ, Swarnakar A;
PI Hafalia AJA, Lindquist EA, Jiang X, Jackson AA, Wilson AD, Jin P;
PI Khare R, Marquis JP;
XX
XX WPI; 2003-269319/26.
DR
DR N-PSDB; ABX12019.
XX

XX Novel human transporter and ion channel polypeptides and polynucleotides
PT for diagnosing, preventing or treating cell proliferative, transport,
PT neurological, muscle and immunological disorders.

XX -Claim 1; Page 216-217; 253pp; English.

XX The invention discloses isolated polypeptides chosen from human
CC transporter and ion channel polypeptides, TRICH 1-26, a biologically
CC active or immunogenic fragment and the nucleic acids encoding them. Also
CC disclosed are isolated antibodies raised against the TRICH proteins,
CC methods for detecting a target polynucleotide in a sample and a
CC microarray where at least one element is a TRICH polynucleotide. The
CC proteins are useful for screening for agonists or antagonists, which can
CC then be used for treating a disease or condition associated with
CC decreased or overexpression of functional TRICH in a patient, for
CC screening for a compound that modulates the activity of the polypeptide
CC or that binds to the polypeptide or as an immunogen for preparing
CC antibodies. The polynucleotides are useful for screening for compounds
CC which alter expression of a target polynucleotide or for assessing
CC toxicity of a test compound. The polypeptides, polynucleotides,
CC modulators and antibodies are useful for diagnosis, treatment (e.g. gene
CC therapy) and prevention of cell proliferative, transport, neurological,
CC muscle and immunological disorders, such as amyotrophic lateral
CC sclerosis, cystic fibrosis, diabetes, Parkinson's disease, prostate
CC cancer, cardiac disorders, angina, Alzheimer's disease, amnesia,
CC epilepsy, schizophrenia, sickle cell anaemia, infertility,
CC hyperglycaemia, hypoglycaemia, hypercholesterolaemia, stroke, multiple
CC sclerosis, motor neuron disorder, prion disease, metabolic disease of the
CC nervous system, developmental disorders of the central nervous system,
CC cardiomyopathy, hypertension, asthma, AIDS, allergies, anaemia,

CC atherosclerosis, atopic dermatitis, diabetes mellitus, osteoarthritis,
CC osteoporosis, rheumatoid arthritis, psoriasis, infections, trauma,
CC hepatitis and cancers, including leukemia and lymphoma. The sequences
CC presented in ABG75819-ABG75844 are the TRICH proteins of the invention
XX
SQ Sequence 507 AA;

Query Match 99.9%; Score 2466; DB 6; Length 507;
Best Local Similarity 99.8%; Pred. No. 1.7e-257;
Matches 475; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTQRLRNEDYHDYSSSTDVSPSESGNLNLSPPGSGYQRFQSGNSTFWQTLHLKGN 60
DB :|||||
32 LSTQRLRNEDYHDYSSSTDVSPSESGNLNLSPPGSGYQRFQSGNSTFWQTLHLKGN 91
QY 61 IGTGLLGLPLAVKXAGIWMGPISLLIIGIVAVHCNGILVKCAHFCRRLNKSFVDYGDIV 120
DB :|||||
92 IGTGLLGLPLAVKXAGIWMGPISLLIIGIVAVHCNGILVKCAHFCRRLNKSFVDYGDIV 151
QY 121 MYGLESPSCWLRNHAHWRVDFFLIVTQLGFCVYFVLADNFKQVIEAANGTTNNC 180
DB :|||||
152 MYGLESPSCWLRNHAHWRVDFFLIVTQLGFCVYFVLADNFKQVIEAANGTTNNC 211
QY 181 HNNETVILTPMTDSRLYMLSELPFLVLLVFNRLPALSIFSLLANITMLVSLWMIYQFIV 240
DB :|||||
212 HNNETVILTPMTDSRLYMLSELPFLVLLVFNRLPALSIFSLLANITMLVSLWMIYQFIV 271
QY 241 QRIIDPDSHLPLVAPWKTYPLFFGTAFSFGEGIMVLPLENKMKDKPRKPLIYLGMVIVT 300
DB :|||||
272 QRIIDPDSHLPLVAPWKTYPLFFGTAFSFGEGIMVLPLENKMKDKPRKPLIYLGMVIVT 331
QY 301 ILXISLGLCYLQFGANIQGSIITLNLPCWLYQSVKLLYSIGIFFTYALQFVPAEIIIP 360
DB :|||||
332 ILXISLGLCYLQFGANIQGSIITLNLPCWLYQSVKLLYSIGIFFTYALQFVPAEIIIP 391
QY 361 FFVSRAPHCBLVDLDFVTVLVCLTCLAILIPLRLDIVSLVGSVSSSALALIIPPLLE 420
DB :|||||
392 FFVSRAPHCBLVDLDFVTVLVCLTCLAILIPLRLDIVSLVGSVSSSALALIIPPLLE 451
QY 421 VTFYSEGMSPLTIFPKDALISILGFGVGVGTYEALYELIOPSNAPIFINSTCAFI 476
DB :|||||
452 VTFYSEGMSPLTIFPKDALISILGFGVGVGTYEALYELIOPSNAPIFINSTCAFI 507

RESULT 4
ABP96441
ID ABP96441 standard; protein; 476 AA.
XX
XX AC ABP96441;
XX
XX DT 02-JUN-2003 (first entry)
XX
XX DE Human tranndorin 3 protein SEQ ID NO:21.
KW Neuroprotective; nootropic; cerebroprotective; analgesic; gene therapy;
KW central nervous system disorder; CNS disorder; multiple sclerosis;
KW nerve injury; neuropathic pain; stroke; trauma; non-CNS disorder; tramdi;
KW tranndorin; human; tramdorin 3; chromosome 5.
XX Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX Misc-difference 356 /note= "encoded by GAG"
XX
XX PN WO2003016502-A2.
XX
XX PD 27-FEB-2003.
XX
XX PF 21-AUG-2002; 2002WO-US026637.
XX
XX PR 21-AUG-2001; 2001US-0313907P.
XX
XX PR 21-AUG-2002; 2002US-00225810.

XX PA	(MCLA-) MCLAUGHLIN RES INST.	DT	04-MAR-2003 (first entry)
XX PI	Birmingham JR;	XX	Human NOVX protein, NOV7b.
XX PI		XX	Human; immunogen; NOVX; metabolic disorder; diabetes; cardiomyopathy;
XX PI		XX	obesity; infectious disease; anorexia; neurodegenerative disorder;
DR DR	WPI; 2003-278567/27.	KW	Alzheimer's disease; Parkinson's disease; immune disorder;
DR DR	N-PSDB; ABZ80236.	KW	haematopoietic disorder; dyslipidaemia; metabolic disturbance;
XX PT	New nucleic acid sequence encoding tramdorins, e.g. mouse tramd 1, mouse	XX	metabolic syndrome X; wasting disorder; cancer; gene therapy.
XX PT	tramd 2, mouse tramd 3, human tramd 1, human tramd 2, human tramd 3 or	OS	Homo sapiens.
XX PT	rat tramd 1, useful for treating CNS, e.g. stroke, multiple sclerosis,	XX	
XX PT	trauma, neuropathic pain.	XX	
XX PS		PN	WO200281518-A2.
XX PS		XX	17-OCT-2002.
XX CC	Example 11; Fig 21; 177pp; English.	XX	
XX CC	The present invention describes an isolated nucleic acid sequence	XX	
XX CC	comprising a cDNA sequence encoding mouse tramdorin (tramd) 2, mouse	XX	
XX CC	tramd 3, human tramd 1, human tramd 2, human tramd 3 or rat tramd 1, or	XX	
XX CC	the genomic sequence of mouse tramd 1 or mouse tramd 3. Mouse tramd 1 is	XX	
XX CC	located to chromosome 11, whereas human tramd 1 is located to chromosome	XX	
XX CC	5q31-33. The tramd sequences have neuroprotective, nootropic, analgesic	XX	
XX CC	and cerebroprotective activities, and can be used in gene therapy. The	XX	
XX CC	nucleic acid sequences are useful for diagnosing and treating central	XX	
XX CC	nervous system (CNS) disorders such as multiple sclerosis, nerve injury,	XX	
XX CC	neuropathic pain, stroke or trauma, and non-CNS disorders. The present	XX	
XX CC	invention represents human tramd 3, which is given in the exemplification	XX	
XX CC	of the present invention	XX	
XX SQ	Sequence 476 AA;	XX	
Query Match		XX	(CURA-) CURAGEN CORP.
Best Local Similarity 99.7%; Score 2461; DB 6; Length 476;		XX	
Matches 475; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		XX	
QY	1 MSTORLNEDYHDYSSTDVSPESPSGLNLSPGSGYQRFQSGNSTTWFQTLHLKGN 60	PI	Pena CEA, Shinkets RA, Li L, Shenoy SG, Kekuda R, Spytek KA;
DB	1 MSTORLNEDYHDYSSTDVSPESPSGLNLSPGSGYQRFQSGNSTTWFQTLHLKGN 60	PI	Vernet CAM, Malyankar UL, Guo X, Gusev VY, Casman SU, Boidog FL;
QY	61 IGTGLGLPLAVKNAGIVMGPIISLLIIGIVAVHCHMGILVKCAHHFCRRLNKSFDYDGTV 120	PI	Furtak K, Tcherev VT, Patturajan M, Gangolli EA, Padigaru M, Liu X;
DB	61 IGTGLGLPLAVKNAGIVMGPIISLLIIGIVAVHCHMGILVKCAHHFCRRLNKSFDYDGTV 120	PI	Baumgartner JC, Gerlach VL, Spaderna SK, Zerhusen BD;
QY	121 MYGLESSPCSWLRNHAHWGRVDFLLIVTQLGFCVCYFVFLADNFKQVIEAANGTTNNC 180	XX	WPI; 2003-046859/04.
DB	121 MYGLESSPCSWLRNHAHWGRVDFLLIVTQLGFCVCYFVFLADNFKQVIEAANGTTNNC 180	XX	N-PSDB; ABX70643.
QY	181 HNNETVILTPMDSRLYMLGPLPLVLLVFRNLRLALSIFSLANITMLVSLVMYQFIV 240	PT	New isolated NOVX polypeptide useful for treating cardiomyopathy,
DB	181 HNNETVILTPMDSRLYMLGPLPLVLLVFRNLRLALSIFSLANITMLVSLVMYQFIV 240	PT	atherosclerosis, metabolic disorders, diabetes, obesity, infectious
QY	241 QRIPDPSHLPLVAPWKTYPIFFGTAFISFEGIGMVLPLENKMKDPKRFPLILVIGMVI 300	PT	disease, anorexia, neurodegenerative disorders, Alzheimer's disease and
DB	241 QRIPDPSHLPLVAPWKTYPIFFGTAFISFEGIGMVLPLENKMKDPKRFPLILVIGMVI 300	PT	cancer.
QY	301 ILYISLCLGYLQPGANIQSITLNPNCMLYQSVKLLYSIGIFFTVALQFYVPAEIIIP 360	XX	Claim 1; Page 57; 479pp; English.
DB	301 ILYISLCLGYLQPGANIQSITLNPNCMLYQSVKLLYSIGIFFTVALQFYVPAEIIIP 360	XX	
QY	361 FFVSRAPHELCVLVDLVRVTVLCTILAILPRLDLIVLSGVSSSALALIIPPLE 420	XX	
DB	361 FFVSRAPHELCVLVDLVRVTVLCTILAILPRLDLIVLSGVSSSALALIIPPLE 420	XX	
QY	421 VTFYSGMSPLTIFKDALISILGFGVFGVTYALYELIQPSNAPFINSTCAFI 476	XX	
DB	421 VTFYSGMSPLTIFKDALISILGFGVFGVTYALYELIQPSNAPFINSTCAFI 476	XX	
RESULT 5		XX	
ABU52588		XX	
ID ABU52588 standard; protein; 496 AA.		XX	
XX		XX	
AC ABU52588;		XX	
XX		XX	

CC assays, detection assays, predictive medicine, and in methods of
CC treatment. NOVX is useful as immunogen, to screen for potential
CC ant/agonist compounds, and as bait protein in a two-hybrid or three-
CC hybrid assay. NOVX NA is useful in gene therapy, to express NOVX, to
CC detect NOVX mRNA or a genetic lesion in a NOVX gene, and to modulate NOVX
CC activity. The cell is useful for producing non-human transgenic animals.
CC Ab is useful for isolating, and purifying NOVX and to monitor protein
CC levels in tissue as part of a clinical testing procedure. The present
CC sequence represents a NOVX protein
xx
SQ Sequence 496 AA:

Sequence 496 AA;

Query Match 99.6%: Score 2460: DB 6: Length 496;

Best Local Similarity 99.8%; Pred. No. 7.4e-257;

Local similarity	Conservative	Mismatches	Indels	Gaps
Matches 475;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;

Qy	1	MSTORLRNEDYHDYSSSTDVSPEESPSEGLNNISPPGYSQRFQGSNSTTWFQTLIHLKGN	60
Db	21	MSTORLRNEDYHDYSSSTDVSPEESPSEGLNNISPPGYSQRFQGSNSTTWFQTLIHLKGN	80
Qy	61	IGTGLLGLPLAVKNAGIWMGPI SLLIIIGIVAVHCMGILVVKCAHHKFCRELKNSFVDYGDTV	120
Db	81	IGTGLLGLPLAVKNAGIWMGPI SLLIIIGIVAVHCMGILVVKCAHHKFCRELKNSFVDYGDTV	140
Qy	121	MYGLESPSCSWLRNHAHWGRVDPFLIVTQLGFCCTVYFVLADNFKQVIEAANGTTNNC	180
Db	141	MYGLESPSCSWLRNHAHWGRVDPFLIVTQLGFCCTVYFVLADNFKQVIEAANGTTNNC	200
Qy	181	HNNETVILTPMTDSRLYMLSFPLFVLVLAVFRNLRALSI FSLLANITMLVSLVLMYIQFIV	240
Db	201	HNNETVILTPMTDSRLYMLSFPLFVLVLAVFRNLRALSI FSLLANITMLVSLVLMYIQFIV	260
Qy	241	QRIPDPSHLPLVAPWKTYPLFFGTAFISFEGIGMVLPLENKKMKOPRKPLLIYLGWIVT	300
Db	261	QRIPDPSHLPLVAPWKTYPLFFGTAFISFEGIGMVLPLENKKMKOPRKPLLIYLGWIVT	320
Qy	301	ILYISLGLGLYQFGANTQGSITLNLPCNLWYQSQVKLLYSIGIFFTYALQFVVPRAEIIIP	360
Db	321	ILYISLGLGLYQFGANTQGSITLNLPCNLWYQSQVKLLYSIGIFFTYALQFVVPRAEIIIP	380
Qy	361	FFVSRAPEHCELVVDLFRVTVLVCLTCILAILPRLDLVISLVGVSSSALALII PPLLE	420
Db	381	FFVSRAPEHCELVVDLFRVTVLVCLTCILAILPRLDLVISLVGVSSSALALII PPLLE	440
Qy	421	VTTFYSEGMSPLTIFKDALISILGFVGVGTYEALYELIQPSNAPIFINSTCAPI	476
Db	441	VTTFYSEGMSPLTIFKDALISILGFVGVGTYEALYELIQPSNAPIFINSTCAPI	496

RESULT 6

AAO14191

ID AAO14191 standard: protein: 568 AA.

XX
XX

AC AAO14191;

XX

DT 03-MAY-2002 (first entry)

XX

DE Human transporter and ion channel TRICH-8.

XX

KW Human; transp

KW neurological

KW cell proliferation:

KW
cerebroprote

gene therapy.

✕ 5

07-JUL-2000; 2000US-0216547P.
14-JUL-2000; 2000US-0218232P.
21-JUL-2000; 2000US-0220112P.
28-JUL-2000; 2000US-0221839P.

(INCY-) INCYTE GENOMICS INC.

Raumann BE, Thornton M, Ding L, Yue H, Tang YT, Harland L;
Burford N, Greene BD, Sanjanwala MS, Baughn MR, Yao MG, Yang J;
Patterson C, Gandhi AR, Hafalia AJA, Tribouley CM, Walia NK;
Pi Au-Young J, Walsh RT, Ramkumar J, Lu Y, Lu DAM, Azimzai Y, Lal P;
Pi Elliott VS, Nguyen DB, Xu Y, Seilhamer JJ, Borowsky ML, Khan FA;
Pi Kearney L, Thangavelu K, Das D, Policky JL;

WPI; 2002-205969/26.
N-PSDB; AAL44674.

New human transporters and ion channel polypeptides for diagnosing,
treating or preventing transport, neurological; muscle, immunological and
cell proliferative disorders.

Claim 1; Page 155-156; 230pp; English.

The present invention provides the protein and coding sequences of a
number of human transporter and ion channel proteins, designated TRICH-1-
TRICH-32. The sequences can be used in the treatment of transport,
neurological, muscle, immunological and cell proliferative disorders. The
present sequence is a protein of the invention

Sequence 568 AA;
SQ

Query Match 97.3%; Score 2402.5; DB 5; Length 568;

Best Local Similarity 88.3%; Pred. No. 1.5e-250;

Sequence	Matches	Mismatches	Indels	Gaps
1	474	0	2	61
2	474	0	2	61
3	474	0	2	61
4	474	0	2	61
5	474	0	2	61
6	474	0	2	61
7	474	0	2	61
8	474	0	2	61
9	474	0	2	61
10	474	0	2	61
11	474	0	2	61
12	474	0	2	61
13	474	0	2	61
14	474	0	2	61
15	474	0	2	61
16	474	0	2	61
17	474	0	2	61
18	474	0	2	61
19	474	0	2	61
20	474	0	2	61
21	474	0	2	61
22	474	0	2	61
23	474	0	2	61
24	474	0	2	61
25	474	0	2	61
26	474	0	2	61
27	474	0	2	61
28	474	0	2	61
29	474	0	2	61
30	474	0	2	61
31	474	0	2	61
32	474	0	2	61
33	474	0	2	61
34	474	0	2	61
35	474	0	2	61
36	474	0	2	61
37	474	0	2	61
38	474	0	2	61
39	474	0	2	61
40	474	0	2	61
41	474	0	2	61
42	474	0	2	61
43	474	0	2	61
44	474	0	2	61
45	474	0	2	61
46	474	0	2	61
47	474	0	2	61
48	474	0	2	61
49	474	0	2	61
50	474	0	2	61
51	474	0	2	61
52	474	0	2	61
53	474	0	2	61
54	474	0	2	61
55	474	0	2	61
56	474	0	2	61
57	474	0	2	61
58	474	0	2	61
59	474	0	2	61
60	474	0	2	61
61	474	0	2	61
62	474	0	2	61
63	474	0	2	61
64	474	0	2	61
65	474	0	2	61
66	474	0	2	61
67	474	0	2	61
68	474	0	2	61
69	474	0	2	61
70	474	0	2	61
71	474	0	2	61
72	474	0	2	61
73	474	0	2	61
74	474	0	2	61
75	474	0	2	61
76	474	0	2	61
77	474	0	2	61
78	474	0	2	61
79	474	0	2	61
80	474	0	2	61
81	474	0	2	61
82	474	0	2	61
83	474	0	2	61
84	474	0	2	61
85	474	0	2	61
86	474	0	2	61
87	474	0	2	61
88	474	0	2	61

Qy	1	MSTQRLNEDYHDYSS	TDVPEESPSGLNNLS	SPGSVQRE	QNS	TTFOTL	HL	LKGN	60
Db	32	MSTQRLNEDYHDYSS	TDVPEESPSGLNNLS	SPGSVQRE	QNS	TTFOTL	HL	LKGN	91
Qy	61	IGTGLLGLPLAVKNAG	IVMGPI	SLLI	IGI	IVAVH	CMGILVK	CAH	HF
Db	92	IGTGLLGLPLAVKNAG	IVMGPI	SLLI	IGI	IVAVH	CMGILVK	CAH	HF
Qy	121	MYGLESPCSWLNRH	HAHGRV	VDFF	LI	VTOLG	FCCVYFV	FLADN	FKO
Db	152	MYGLESPCSWLNRH	HAHGRV	VDFF	LI	VTOLG	FCCVYFV	FLADN	FKO
Qy	181	HNNETVILPTMDSRL	YMWLSF	LP	LVLL	VFTRNL	RALS	I	F
Db	212	HNNETVILPTMDSRL	YMWLSF	LP	LVLL	VFTRNL	RALS	I	F
Qy	241	QRIPDPGHLPLVAP	PKWTY	PL	FFG	TAI	FS	FEG	IG
Db	272	QRIPDPGHLPLVAP	PKWTY	PL	FFG	TAI	FS	FEG	IG
Qy	301	ILYISLCGLYLQF	GANI	QGS	ITL	N	PNC	WLY	Q
Db	332	ILYISLCGLYLQF	GANI	QGS	ITL	N	PNC	WLY	Q
Qy	361	FFVSRAPCEHCELV	VDLF	VRT	VL	VCIT	-----	-----	386
Db	392	FFVSRAPCEHCELV	VDLF	VRT	VL	VCIT	-----	-----	451
Qy	387	-----C-----	-----	-----	-----	-----	-----	-----	419
Db	452	HPWLSPFCEQWVS	CHRD	T	VV	WMGF	ARG	I	LA
Qy	420	EVTTFFYSEGMS	PLIF	K	D	A	L	I	S
Db	512	EVTTFFYSEGMS	PLIF	K	D	A	L	I	S

RESULT 7

ABU52587
ID ABU52587 standard; protein; 500 AA.
XX
AC ABU52587;
DT
XX 04-MAR-2003 (first entry)
DE Human NOVX protein, NOV7a.
XX
KW Human; immunogen; NOVX; metabolic disorder; diabetes; cardiomyopathy;
KW obesity; infectious disease; anorexia; neurodegenerative disorder;
KW Alzheimer's disease; Parkinson's disease; immune disorder;
KW haematopoietic disorder; dyslipidaemia; metabolic disturbance;
KW metabolic syndrome X; wasting disorder; cancer; gene therapy; SNP;
KW single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 173
FT /note= "May be Val as a result of a single nucleotide
FT polymorphism"
XX
XX WO200281518-A2.
XX
XX *17-OCT-2002.
XX
XX 21-FEB-2002; 2002WO-US005374.
XX
XX 21-FEB-2001; 2001US-0270220P.
XX 21-FEB-2001; 2001US-0270523P.
XX 23-FEB-2001; 2001US-0270797P.
XX 23-FEB-2001; 2001US-0270810P.
XX 08-MAR-2001; 2001US-0274295P.
XX 16-MAR-2001; 2001US-0276400P.
XX 16-MAR-2001; 2001US-0276677P.
XX 26-MAR-2001; 2001US-0278796P.
XX 04-APR-2001; 2001US-0281521P.
XX 25-APR-2001; 2001US-0286548P.
XX 17-MAY-2001; 2001US-0291765P.
XX 10-AUG-2001; 2001US-0311595P.
XX 10-AUG-2001; 2001US-0311980P.
XX 10-SEP-2001; 2001US-0318526P.
XX 17-SEP-2001; 2001US-0322712P.
XX 18-OCT-2001; 2001US-0330307P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Pena CEA, Shinkets RA, Li L, Shenoy SG, Kekuda R, Spytek KA;
XX Vernet CAM, Malyankar UL, Guo X, Gusev VY, Casman SJ, Boidog FL;
XX Furtak K, Tchernev VT, Patturajan M, Gangolli EA, Padigar M, Liu X;
XX Baumgartner JC, Gerlach VL, Spaderna SK, Zerhusen BD;
XX
XX WPI; 2003-046859/04.
XX N-PSDB; ABX70642.
XX
XX New isolated NOVX polypeptide useful for treating cardiomyopathy,
XX atherosclerosis, metabolic disorders, diabetes, obesity, infectious
XX disease, anorexia, neurodegenerative disorders, Alzheimer's disease and
XX cancer.
XX
XX Claim 1; Page 56; 479pp; English.
XX
XX The invention relates to an isolated polypeptide termed NOVX (NOV1, 2a,
XX 2b, 3a, 3b, 4a, 4b, 5, 6, 7a-c, 8a-e, 9a-b, 10, 11, 12a-c, 13, 14, 15,
XX 16a-d, 17a-b, 18, 19, 20a-b, 21-30) appearing as ABU52578-ABU52624), a
XX variant of NOVX, a mature form of NOVX, and a variant of the mature form
XX of NOVX. Also included are a nucleic acid molecule (NOVX NA) encoding
XX NOVX, or a fragment or complement of NOVX NA, a vector comprising NOVX
XX NA, a cell comprising the vector, an anti-NOVX antibody (ab), determining
XX the presence or amount of NOVX or NOVX NA in a sample, and identifying an
XX agent that binds or modulates the expression or activity of NOVX. NOVX,
XX NOVX NA or ab is useful for treating or preventing a NOVX-associated

disorder in a subject, preferably human. Ab is useful for determining the presence or amount of NOVX in a sample. NOVX is useful for identifying an agent that binds to NOVX. NOVX, NOVX NA or ab is useful for treating metabolic disorders, diabetes, cardiomyopathy, obesity, infectious disease, anorexia, neurodegenerative disorders, Alzheimer's disease, Parkinson's disease, immune disorders, haematopoietic disorders, and various dyslipidaemias, metabolic disturbances associated with obesity, the metabolic syndrome X and wasting disorders associated with chronic diseases, various cancers, endocrine, connective tissue, blood, vascular, skin, renal, bone, brain, muscle disorders, or bacterial, fungal, protozoal or viral infections. NOVX, NOVX NA or ab is useful in screening assays, detection assays, predictive medicine, and in methods of treatment. NOVX is useful as immunogen, to screen for potential ant/agonist compounds, and as bait protein in a two-hybrid or three-hybrid assay. NOVX NA is useful in gene therapy, to express NOVX, to detect NOVX mRNA or a genetic lesion in a NOVX gene, and to modulate NOVX activity. The cell is useful for producing non-human transgenic animals. Ab is useful for isolating, and purifying NOVX and to monitor protein levels in tissue as part of a clinical testing procedure. The present sequence represents a NOVX protein

Query Match 97.1%; Score 2398; DB 6; Length 500;
Best Local Similarity 97.5%; Pred. No. 3.9e-250;
Matches 472; Conservative 0; Mismatches 0; Indels 12; Gaps 3;

QY 1 MSTQRLRNEDYHDYSDTSPESPSSEGLNLSPGSYQRFQGSNSTTWFTQLHLKGN 60
DB 21 MSTQRLRNEDYHDYSDTSPESPSSEGLNLSPGSYQRFQGSNSTTWFTQLHLKGN 80
QY 61 IGTGLLGLPLAVKNAGIVMGPIISLLIIGIVAVHGMGILVKCAHFCRRLNKSFVDYGDV 120
DB 81 IGTGLLGLPLAVKNAGIVMGPIISLLIIGIVAVHGMGILVKCAHFCRRLNKSFVDYGDV 140
QY 121 MYGLESPCSWLRNHAHWGRVVDFFLIYVTLQGLGCCVYFVFLADNFKQVIEAANGTTNCC 180
DB 141 MYGLESPCSWLRNHAHWGRVVDFFLIYVTLQGLGCCVYFVFLADNFKQVIEAANGTTNCC 200
QY 181 HNNETVILPTMDSRLYMLSFPLVLLVFNRLRLAISIFSLLANITMLVSLVMYQFIV 240
DB 201 HNNETVILPTMDSRLYMLSFPLVLLVFNRLRLAISIFSLLANITMLVSLVMYQFIV 260
QY 241 -----QRIPDPHPLVAPWKTYPLFGTTFPSFEGIGMVLPLENKWKDPRKFPPLIL 292
DB 261 FRYMLSVFQRIQIPDPHPLVAPWKTYPLFGTTFPSFEGIGMVLPLENKWKDPRKFPPLIL 320
QY 293 YLGMVITVILYISLGLGVLOFGANTIGSITNLNPNCLWYQSVKLYSIGIFFTYALQFY 352
DB 321 YLGMVITVILYISLGLGVLOFGANTIGSITNLNPNCLWYQSVKLYSIGIFFTYALQFY 377
QY 353 VPAEIIIPFVSRAPEHCELVVDLFRVTVLVCLTCLAILIPRLDLVISLVGSSSALA 412
DB 378 VPAEIIIPFVSRAPEHCELVVDLFRVTVLVCLTCLAILIPRLDLVISLVGSSSALA 436
QY 413 LIIPPLEVTTFYSEGMSPLTFKDALISILGFVGVFGVYEALEYLIQPSNAPIFINST 472
DB 437 LIIPPLEVTTFYSEGMSPLTFKDALISILGFVGVFGVYEALEYLIQPSNAPIFINST 496
QY 473 CAFE 476
DB 497 CAFE 500

RESULT 8
ABU76941
ID ABU76941 standard; protein; 476 AA.
XX
AC ABU76941;
XX
XX 22-JUL-2002 (first entry)
XX
XX Human VG51.

XX Antiasthmatic; anxiolytic; antiepileptic; antihypertensive; human;
KW psychotropic; glutamate transporter; transporter; GABA;
KW gamma-aminobutyric acid transporter; GABA transporter; neurotransmitter;
KW asthma; anxiety; epilepsy; hypertension; psychiatric disorder;
KW neurotic disorder; VG51.
XX
OS Homo sapiens.
XX
XX WO200071709-A1.
XX
XX 30-NOV-2000.
XX
XX 19-MAY-2000; 2000WO-FR001383.
XX
XX 21-MAY-1999; 99FR-00006525.
XX
XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX
XX Giros B, Gasnier B, Sagne C, El Mestikawy S, Hamon M;
XX
XX WPI; 2001-025160/03.
XX
XX N-PSDB; ABL57930.
XX
XX New mammalian amino acid transporter, used e.g. to screen for
XX psychotropic agents, is high capacity but low affinity transporter of
XX gamma-aminobutyric acid.
XX
XX Claim 2; Fig 13; 103pp; French.
XX
XX The present sequence is the protein sequence for human VG51, a glutamate/
XX gamma-aminobutyric acid (GABA) transporter. GABA and glutamate are
XX neurotransmitters. The transporter can be used to produce specific
XX antibodies, to screen for binding agents. Modulators of the transporter
XX are useful for treating disorders associated with deregulated
XX glutamate/GABA transport, e.g. asthma, anxiety, epilepsy, hypertension
XX and other psychiatric and neurotic disorders, while determining levels of
XX the transporter and its coding sequence can be used for diagnosis of such
XX disorders
XX
XX Sequence 476 AA;
SQ
Query Match 94.1%; Score 2324; DB 4; Length 476;
Best Local Similarity 94.5%; Pred. No. 3.8e-242;
Matches 450; Conservative 9; Mismatches 17; Indels 0; Gaps 0;
Qy 1 MSTQRLNEDYHDSSTDVSPSESGNLNLSFGSYQRFQGSNSTTWFQTLIHLKGN 60
Db 1 MSTQRLNEDYHDSSTDVSPSESGNLNLSFGSYQRFQGSNSTTWFQTLIHLKGN 60
Qy 61 IGTGLGLPLAVKAGIWMGPISLLIIGIVAVHGMGILVKAHFCRLNKSFDYDGTIV 120
Db 61 IGTGLGLPLAVKAGIWMGPISLLIIGIVAVHGMGILVKAHFCRLNKSFDYDGTIV 120
Qy 121 MYGLESPCSWLNRHAGRWVDFLLVITQLGFCVYFVFLADNFKQVIEAANGTTNNC 180
Db 121 MYGLESPCSWLNRHAGRWVDFLLVITQLGFCVYFVFLADNFKQVIEAANGTTNNC 180
Qy 181 HNNETVILPTMTDSRLYMLSPFLVLVLFIRNLRALSIFSLLANITMLVSLVMYQFIV 240
Db 181 HNNETVILPTMTDSRLYMLSPFLVLVLFIRNLRALSIFSLLANITMLVSLVMYQFIV 240
Qy 241 QRIIDPSHLPLVAPWKTYPLEFGTAIFSEFEGIMVLPLENKMDPRKPEPLIYLGMVIT 300
Db 241 QRIIDPSHLPLVAPWKTYPLEFGTAIFSEFEGIMVLPLENKMDPRKPEPLIYLGMVIT 300
Qy 301 ILYISLGLCYLQFGANIQSGITLNLPCWLYQSVKLYSGIFFTYALQFVVAEIIIP 360
Db 301 ILYISLGLCYLQFGANIQSGITLNLPCWLYQSVKLYSGIFFTYALQFVVAEIIIP 360
Qy 361 FFVSRAPHCHELVDFLVRTVLVCUTCTLAILIPRLDLVLSVGSVSSALALIIPPLE 420
Db 361 VIVSWCKCTLMVDLIGSAMLCKTCLAILIPRLDLVLSVGSVSSALALIIPPLE 420

Qy 421 VTTYSEGMSPLTIFKDALISILGVGVVGTVEALYELIQPSNAPINFINSTCAFI 476
Db 421 VTTYSEGMSPLTIFKDALISILGVGVVGTVEALYELIQPSNAPINFINSTCAFI 476
RESULT 9
ADF90169
ID ADF90169 standard; protein; 476 AA.
XX
XX ADF90169;
XX
XX 26-FEB-2004 (first entry)
XX
XX Human transporter homologue seq id 7.
XX
XX peptide therapy; transporter; human.
XX
XX Homo sapiens.
XX
XX US2003170778-A1.
XX
XX 11-SEP-2003.
XX
XX 14-MAR-2001; 2001US-00805456.
XX
XX 22-DEC-2000; 2000US-0257175P.
XX
XX (WEIM/) WEI M.
XX (YANC/) YAN C.
XX (MERK/) MERKLOV G V.
XX (KETC/) KETCHUM K A.
XX (DFRA/) DI FRANCESCO V.
XX (BEAS/) BEASLEY E M.
XX
XX Wei M, Yan C, Merklov GV, Ketchum KA, Di Francesco V, Beasley EM;
XX WPI; 2003-898259/82.
XX
XX New human transporter peptide, useful for preparing a composition for
XX treating a disease or condition mediated by human transporters.
XX
XX Disclosure; SEQ ID NO 7; 94pp; English.
XX
XX The invention describes an isolated peptide comprising a 476 amino acid
XX sequence (S1) or its allelic variant, orthologue or fragment, where the
XX allelic variant or orthologue is encoded by a nucleic acid that
XX hybridises under stringent conditions to the opposite strand of a nucleic
XX acid comprising a sequence having 2093 or 46649 base pairs, and the
XX fragment comprises 10 contiguous amino acids of S1. The peptide is useful
XX for preparing a composition for treating a disease or condition mediated
XX by a human transporter protein. This is the amino acid sequence of a
XX human transporter protein homologue.
XX
XX Sequence 476 AA;
SQ
Query Match 94.1%; Score 2324; DB 7; Length 476;
Best Local Similarity 94.5%; Pred. No. 3.8e-242;
Matches 450; Conservative 9; Mismatches 17; Indels 0; Gaps 0;
Qy 1 MSTQRLNEDYHDSSTDVSPSESGNLNLSFGSYQRFQGSNSTTWFQTLIHLKGN 60
Db 1 MSTQRLNEDYHDSSTDVSPSESGNLNLSFGSYQRFQGSNSTTWFQTLIHLKGN 60
Qy 61 IGTGLGLPLAVKAGIWMGPISLLIIGIVAVHGMGILVKAHFCRLNKSFDYDGTIV 120
Db 61 IGTGLGLPLAVKAGIWMGPISLLIIGIVAVHGMGILVKAHFCRLNKSFDYDGTIV 120
Qy 121 MYGLESPCSWLNRHAGRWVDFLLVITQLGFCVYFVFLADNFKQVIEAANGTTNNC 180
Db 121 MYGLESPCSWLNRHAGRWVDFLLVITQLGFCVYFVFLADNFKQVIEAANGTTNNC 180
Qy 181 HNNETVILPTMTDSRLYMLSPFLVLVLFIRNLRALSIFSLLANITMLVSLVMYQFIV 240
Db 181 HNNETVILPTMTDSRLYMLSPFLVLVLFIRNLRALSIFSLLANITMLVSLVMYQFIV 240

Db	181	HNNETVILPTMTDSRLYMLSFPLVLLVFIRNLRLSIFSLLANITMLVSLVMIYQFIV	240
Qy	241	QRIIPDPSHLPLVAPWKTYPLFFGTAIFSPFEGIGWVLPLENKMDPRKFPILYLGMVIVT	300
Db	241	QRIIPDPSHLPLVAPWKTYPLFFGTAIFSPFEGIGWVLPLENKMDPRKFPILYLGMVIVT	300
Qy	301	ILYISLGLGYLQFGANIQGSITLNLPCNWLQSVLLYSGISGIFFTYVALQFYVPAEIIIP	360
Db	301	ILYISLGLGYLQFGANIQGSITLNLPCNWLQSVLLYSGISGIFFTYVALQFYVPAEIIIP	360
Qy	361	FFVSRAPHECELVDLFFVRLVCLTCILAILPRDLVLSLGSVSSSALALIIPPLE	420
Db	361	VIVSWCKCTLMVDLGIGSAMLCKTCILAILPRDLVLSLGSVSSSALALIIPPLE	420
Qy	421	VTFYSEGMSPLTIFKDALISILGFGVGVGTVEALYELIQSNAPIFINSTCAFI	476
Db	421	VTFYSEGMSPLTIFKDALISILGFGVGVGTVEALYELIQSNAPIFINSTCAFI	476
RESULT 10			
ADRO3804			
ID	ADRO3804	standard; protein; 476 AA.	
XX	AC	ADRO3804;	
XX	DT	21-OCT-2004 (first entry)	
XX	DE	Human homologue of novel transporter protein.	
XX	KW	Human; transporter.	
XX	OS	Homo sapiens.	
XX	PN	US2004146887-A1.	
XX	PD	29-JUL-2004.	
XX	PF	07-OCT-2003; 2003US-00679362.	
XX	PR	22-DEC-2000; 2000US-0257175P.	
XX	PR	14-MAR-2001; 2001US-00805456.	
XX	PA	(APPL-) APPLERA CORP.	
XX	PI	Wei M, Yan C, Merklov G, Ketchum KA, Difrancesco V, Beasley EM;	
XX	XX	WPI; 2004-552659/53.	
XX	PT	New nucleic acid encoding human transporter peptides, useful for	
XX	PT	preparing agents for treating a disease or condition mediated by human	
XX	PT	transporters.	
XX	PS	Disclosure; SEQ ID NO 7; 62pp; English.	
XX	CC	The invention relates to a new isolated nucleic acid comprising a	
XX	CC	sequence encoding a transporter protein (appearing as ADR03799) and	
XX	CC	appearing as ADR03798 (the cDNA) or ADR03800 (the gene). Also included	
XX	CC	are a vector comprising the isolated nucleic acid, a host cell containing	
XX	CC	the vector, producing the polypeptide, detecting the presence of the	
XX	CC	nucleic acid in a sample and an isolated human peptide having a sequence	
XX	CC	that shares at least 70% homology with ADR03799. The gene encoding the	
XX	CC	human transporter is located on chromosome 5. The peptide is useful for	
XX	CC	preparing agents for treating a disease or condition mediated by a human	
XX	CC	transporter protein. The present sequence represents a homologue of the	
XX	CC	transporter protein.	
XX	SQ	Sequence 476 AA;	
Query Match		94.1%; Score 2324; DB 8; Length 476;	
Best Local Similarity		94.5%; Pred. No. 3.8e-242;	
Matches 450; Conservative		9; Mismatches 17; Indels 0; Gaps 0;	
Qy	1	MSTQRLRNEDYHDYSSTVSPRESPEGLNNLSSPGSYQRFQGSNSTTWFTLIHLKGN	60
Db	1	MSTQRLRNEDYHDYSSTVSPRESPEGLNNLSSPGSYQRFQGSNSTTWFTLIHLKGN	60
Qy	61	IGTGLLGLPLAVKNAGIVMGPIISLLIIGIVAVHGMGILVKCAHFCRRLNKSFVDYGDIV	120
Db	61	IGTGLLGLPLAVKNAGIVMGPIISLLIIGIVAVHGMGILVKCAHFCRRLNKSFVDYGDIV	120
Qy	121	MYGLESPCSWLNRHAWGRRVVDFELIVTQLGFCVYFVFLADNFKQVIEAANGTTNNC	180
Db	121	MYGLESPCSWLNRHAWGRRVVDFELIVTQLGFCVYFVFLADNFKQVIEAANGTTNNC	180
Qy	181	HNNETVILPTMTDSRLYMLSFPLVLLVFIRNLRLSIFSLLANITMLVSLVMIYQFIV	240
Db	181	HNNETVILPTMTDSRLYMLSFPLVLLVFIRNLRLSIFSLLANITMLVSLVMIYQFIV	240
Qy	241	QRIIPDPSHLPLVAPWKTYPLFFGTAIFSPFEGIGWVLPLENKMDPRKFPILYLGMVIVT	300
Db	241	QRIIPDPSHLPLVAPWKTYPLFFGTAIFSPFEGIGWVLPLENKMDPRKFPILYLGMVIVT	300
Qy	301	ILYISLGLGYLQFGANIQGSITLNLPCNWLQSVLLYSGISGIFFTYVALQFYVPAEIIIP	360
Db	301	ILYISLGLGYLQFGANIQGSITLNLPCNWLQSVLLYSGISGIFFTYVALQFYVPAEIIIP	360
Qy	361	FFVSRAPHECELVDLFFVRLVCLTCILAILPRDLVLSLGSVSSSALALIIPPLE	420
Db	361	VIVSWCKCTLMVDLGIGSAMLCKTCILAILPRDLVLSLGSVSSSALALIIPPLE	420
Qy	421	VTFYSEGMSPLTIFKDALISILGFGVGVGTVEALYELIQSNAPIFINSTCAFI	476
Db	421	VTFYSEGMSPLTIFKDALISILGFGVGVGTVEALYELIQSNAPIFINSTCAFI	476
RESULT 11			
ADAS5508			
ID	ADAS5508	standard; protein; 434 AA.	
XX	AC	ADAS5508;	
XX	DT	20-NOV-2003 (first entry)	
XX	DE	Human protein, SEQ ID 3076.	
XX	KW	Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;	
XX	KW	Gene Therapy; human; secretory protein; membrane proteins; cancer;	
XX	KW	inflammatory disease; osteoporosis; neurological disease.	
XX	OS	Homo sapiens.	
XX	PN	EP1293569-A2.	
XX	PD	19-MAR-2003.	
XX	PF	21-MAR-2002; 2002EP-00006586.	
XX	PR	14-SEP-2001; 2001JP-00328381.	
XX	PR	24-JAN-2002; 2002US-0350435P.	
XX	PA	(HELI-) HELIX RES INST.	
XX	PA	(REAS-) RES ASSOC BIOTECHNOLOGY.	
XX	PI	Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;	
XX	PI	Yamanoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irrie R, Tamechika I;	
XX	PI	Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;	
XX	XX	WPI; 2003-395539/38.	
XX	DR	N-PSDB; ADA53869.	
XX	PT	New polynucleotides encoding full-length polypeptides, e.g. secretory	
XX	PT	and/or membrane proteins, useful for developing medicines for diseases in	
XX	PT	which the gene is involved, or as target molecules for gene therapy.	
XX	XX	Claim 14; SEQ ID NO 3076; 205pp; English.	

nerve injury; neuropathic pain; stroke; trauma; non-CNS disorder; tramd.
tramdorin; rat; LYAAT-1.
XX
OS Rattus sp.
XX
PN W02003016502-A2.
XX
PD 27-FEB-2003.
XX
XX 21-AUG-2002; 2002WO-US026637.
XX
XX 21-AUG-2001; 2001US-0313907P.
PR
XX 21-AUG-2002; 2002US-00225810.
PR
XX
PA (MCLA-) MCLAUGHLIN RES INST.
XX
PI Bermingham JR;
XX
XX WPI; 2003-278567/27.
DR
DR N-PSDB; ABZ80241.
XX
PT New nucleic acid sequence encoding tramdorins, e.g. mouse tramd 1, mouse
tramd 2, mouse tramd 3, human tramd 1, human tramd 2, human tramd 3 or
rat tramd 1, useful for treating CNS, e.g. stroke, multiple sclerosis,
trauma, neuropathic pain.
PT
PT
PS Disclosure; Fig 24; 177pp; English.
XX
XX The present invention describes an isolated nucleic acid sequence
comprising a cDNA sequence encoding mouse tramdorin (tramd) 2, mouse
tramd 3, human tramd 1, mouse tramd 2, human tramd 3 or rat tramd 1, or
the genomic sequence of mouse tramd 1 or mouse tramd 3. Mouse tramd 1 is
located to chromosome 11, whereas human tramd 1 is located to chromosome
5q31-33. The tramd sequences have neuroprotective, nootropic, analgesic
and cerebroprotective activities, and can be used in gene therapy. The
nucleic acid sequences are useful for diagnosing and treating central
nervous system (CNS) disorders such as multiple sclerosis, nerve injury,
neuropathic pain, stroke or trauma, and non-CNS disorders. The present
sequence represents rat LYAAT-1, which is given in the exemplification of
the present invention
XX
SQ Sequence 475 AA;
Query Match 86.9%; Score 2146.5; DB 6; Length 475;
Best Local Similarity 85.7%; Pred No. 6.3e-223;
Matches 408; Conservative 36; Mismatches 31; Indels 1; Gaps 1;
Qy 1 MSTQRLNEDYHDYSTDVSPEESPSEGLNNLSPGSYQRFQSGNSTWFTQLHLKGN 60
Db 1 MSTQRLNEDYHDYSTDVSPEESPSEGLGSP-SPGSYQRLGENSMTWFTQLHLKGN 59
Qy 61 IGTGLGLPLAVKNAAGVIMPTLSLLIIGIVAVHCGMILVKCAHHFCRRLNKSFVDYGDV 120
Db 60 IGTGLGLPLAVKNAAGLGLLSLVIGIVAVHCGMILVKCAHHLCRRLNKPFIDYGDV 119
Qy 121 MYGLESPCSWLNRNHAHGRVVDVFLIVTQLGCCVYFVLADNFKQVTEAAGTTTNC 180
Db 120 MYGLESPSTWIRNHSWGRVVDVFLVVTQLGCCVYFVLADNFKQVTEAAGTTTNC 179
Qy 181 HNNETVILPTMDSRLYMLSFPLVLLVIFRNLRLAISFLLANITMLVSLVMIYQFV 240
Db 180 NNNETVILPTMDSRLYMLTFPLVLLVSFIENLRILSIFLLANISMVFSLMIMYQFV 239
Qy 241 QRIPDPSHLPLVAPWKTYPLPFGTAFISFEGIGMVLPLENKKQDPKRPPLILYLGWIVT 300
Db 240 QRIPDPSHLPLVAPWKTYPLPFGTAFIFAPEGIGVVLPLENKKWQSKQFPPLILYLGMAIT 299
Qy 301 ILVYISGCGYLOFGANIOGSTITLNPNCWLYQSVKLLYSIGIFETVALQFYVPAETIIP 360
Db 300 VLVISGSLGYQFGADIKGSTITLNPNCWLYQSVKLLYSIGIFETVALQFYVAAETIIP 359
Qy 361 FFVSRAPEHCELVVDVFRVTVLVCLTICAILIPRLDLVILSVGSSVSSALALIPPLE 420

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2006, 01:19:59 ; Search time 41 Seconds
(without alignments)
1117.053 Million cell updates/sec

Title: US-10-679-362-2

Perfect score: 2469

Sequence: 1 MSTQRLRNEDYHYSSTDVS.....YELIQSNAPIFINSTCAFI 476

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_80.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1677.5	67.9	481	JC7961	proton-coupled ami
2	672	27.2	607	T26845	hypothetical prote
3	652	26.4	449	H88022	protein T27A1.5 [i
-4	465.5	18.9	436	T05653	amino acid transpo
5	465.5	18.9	713	S58251	probable membrane
6	464	18.8	434	T15799	hypothetical prote
7	462.5	18.7	481	T23131	hypothetical prote
8	443.5	18.0	460	F88544	protein F59B2.2 [i
9	431.5	17.5	467	T26705	hypothetical prote
10	415	16.8	656	T38741	major facilitator
11	388.5	15.7	692	S37976	hypothetical prote
12	368	14.9	389	T31123	hypothetical prote
13	285.5	11.6	543	T48239	hypothetical prote
14	276	11.2	516	T48238	hypothetical prote
15	274.5	11.1	426	T51506	hypothetical prote
16	270.5	11.0	529	E84813	hypothetical prote
17	263	10.7	503	T26524	hypothetical prote
18	259	10.5	890	T21000	hypothetical prote
19	251.5	10.2	462	S42372	hypothetical prote
20	251.5	10.2	486	T42254	amino acid permeas
21	247	10.0	423	T49959	hypothetical prote
22	233.5	9.5	571	T06737	hypothetical prote
23	233	9.4	484	T34016	hypothetical prote
24	232	9.4	509	S45413	probable membrane
25	230	9.3	494	T16658	hypothetical prote
26	224	9.1	476	C96505	probable amino aci
27	222	9.0	505	T88206	protein F21D12.3 [
28	215.5	8.7	485	A48187	amino acid transpo
29	210	8.5	448	S50622	hypothetical prote

ALIGNMENTS

RESULT 1

JC7961

proton-coupled amino acid transporter 2 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 25-Aug-2003 #sequence_revision 25-Aug-2003 #text_change 15-Sep-2003

C:Accession: JC7961

R:Chen, Z.; Kennedy, D.J.; Wake, K.A.; Zhuang, L.; Ganapathy, V.; Thwaites, D.T.

Biochem. Biophys. Res. Commun. 304, 747-754, 2003.

A:Title: Structure, tissue expression pattern, and function of the amino acid transporter

A:Reference number: JC7961; PMID:12727219

A:Accession: JC7961

A:Molecule type: mRNA

A:Residues: 1-481 <CH>

A:Experimental source: (Lung)

C:Comment: This transport protein is a second member of H⁺-coupled, pH-dependent, Na⁺-in

sport not only from one cellular compartment to another (across plasma membranes) but al

C:Genetics:

A:Gene: pat2

A:Map position: 10q22

C:Keywords: amino acid transport system; PAT2

Query Match	67.9%	Score	1677.5	DB 2	Length	481			
Best Local Similarity	71.7%	Pred. No.	2.7e-127						
Matches	325	Conservative	54	Mismatches	67	Indels	7	Gaps	1
QY	20	SPESPSBGLNLS	SPGSYQRFQGSNSTTWFTLHLKGNIGTGLLGLPLAVKNAGIVM	79					
Db	32	SQDPNPVNGSS	ESSE-----KTKGITGFQTLVHLVKGNMGTGILGLPLAVKNAGILM	84					
QY	80	GPISLLIIGIVAVH	CMGILVKCAHFCRLNKSFDYGDVTVMYGLESSPCSWLRNHAHWG	139					
Db	85	GPLSLLVWGLIACH	CMHILVRCARFCHRLNKPMDYGDVTVMHGLASSPNTWLQSHAHWG	144					
QY	140	RRVVDFFLIVTQLG	FCVCYFVELADNFKQVTEAANGTTNNCHNETVILTPMDSRLYML	199					
Db	145	RHAVSFLLIVTQLG	FCVCYIVFLADNLKQVVEAVNSTTISCHKNETVLTPTIDSLRYML	204					
QY	200	SFLPFLVLLVIRNL	RALSIFSLANITMLVSLVMYQFIVQRIQIPDPDPSHLPLVAPWKYTP	259					
Db	205	AFLPVLGLLVIRNL	RVLTIFSLLANVSLVSLVIQYIIQGIQIPDPSPQLPLVASKYTP	264					
QY	260	LFPGTALFSGEGT	WVLPLENKMDPKRFPFLIYLGWIVTILYSLCLGLQFQGANIQ	319					
Db	265	LFPGTALFSGESG	WVLPLENKMDARRFPITLSLGMSTIITLYAIGALGLRFGDDIK	324					
QY	320	GSITLNLPCNLWY	QSVKLLYSIGIFFYVALQFYVPAEIIIPFVSRAPHECHLVVDLVR	379					
Db	325	ASITLNLPCNLWY	QSVKLLYVVGILCTHALQFYVPAEIIIPLAVSQVSKRWALPVDLSIR	384					
QY	380	TVLVCLTCILAIL	IPRLDLVLSVGSVSSSALALIIPPLLEVTTPYSGMGLTTFKQAL	439					
Db	385	LALVCVTCMLAIL	IPRLDLVLSVGSVSSSALALIIPPLLEVTTPYSGMGLTTFKQAL	444					

[illegible]

Db 37 FKTANFVAVGAGVLGLPYAFKRTGMLGVLLSVSVLTHRCMLLV-----YTRRK 91
Qy 110 NKSE-----VDYGDVTVMYGLSSPCSWLRNHAHGRVVDFFLIVTQLGFCVCFVFL 162
Db 92 LDSFNAGISKIGSGD-----LGFVCGSL-----GRIVVDLFIILSQAGFCVGLIFI 140
Qy 163 ADNPKQVIEAANGTT--NNCHNETVILTP-----MDSRLYMLSFPLFL 205
Db 141 -----GTTLANLSDPE-----SPTSLRHQFTRLGSEFLGVSSKSLIYIWGCPFPQ 184
Qy 206 VLLVFIRNLRLSIFSLLANI-----TMLSVAMVYQFIVQRI-PDPShLPVAPKTYPLF 261
Db 185 LGENSIKTLTHLAPLSIFADIVDGAMAVVIVDSMILKORP-----VVAEGGMSLF 238
Qy 262 ---PGTAIFSGEGMVLPLENKKMDPKFPLIILYLGWVITVILYISLGLCYGLQFGANI 318
Db 239 LYNGVAVYSEGVMVLPLESEMKKDKFKGVLAGNGFISLIYIAGFILGYLAFGEDT 298
Qy 319 QGSIITLNPCLWYQSVKLYSIGIFFTYALQFVVPABIIIPFVSVRAPEHCELVDLFFV 378
Db 299 MDIITANLGLAVSTVVLQGLCINLFTFTPLMNPVFEIVERRF--SRG-----MYSAWL 351
Qy 379 RTVLAVCLTCLAILIPLRLDLVISLVGSSSALALIPPLLEVTFYSEGMSPLTIKOA 438
Db 352 RWWLVAVTLVALFVPNFADFLSLVGSSTCCVGLGVLPALPHLLVF--KEEMGWLQWSSDT 410
Qy 439 LISILGFVGVVGYEALYEL 459
Db 411 AIVLVGVVAVSGTWSSLSLSEI 431

RESULT 5
S58251
Probable membrane protein YNL101w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein N2185
C:Species: Saccharomyces cerevisiae
C:Date: 13-Jan-1996 #sequence revision 23-Feb-1996 #text_change 09-Jul-2004
C:Accession: S58251; S63041; S63953
R:Saiz, J.E.; Buitrago, M.J.; Soler, A.; del Rey, F.; Revuelta, J.L.
submitted to the EMBL Data Library, July 1995
A:Description: The sequence of a 21.3 kb fragment from the left arm of yeast chromosome XI
A:Reference number: S58246
A:Accession: S58251
A:Molecule type: DNA
A:Residues: 1-713 <SAI>
A:Cross-references: UNIPROT:P50944; UNIPARC:UPI000005256D; EMBL:Z50161; NID:G929846; PID
R:Saiz, J.E.; Buitrago, M.J.; Soler-Mira, A.; del Rey, F.; Revuelta, J.L.
submitted to the Protein Sequence Database, April 1996
A:Reference number: S63037
A:Accession: S63041
A:Molecule type: DNA
A:Residues: 1-713 <SAW>
A:Cross-references: UNIPARC:UPI000005256D; EMBL:Z71377; NID:gl302013; PIDN:CAA95977.1; F
A:Experimental source: strain S288C
R:Saiz, J.E.; Buitrago, M.J.; Soler-Mira, A.; del Rey, F.; Revuelta, J.L.
Yeast 12, 403-409, 1996
A:Title: The sequence of a 21.3 kb DNA fragment from the left arm of yeast chromosome XI
A:Reference number: S63948; MUID:96267765; PMID:8701612
A:Accession: S63953
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-713 <SAF>
A:Cross-references: UNIPARC:UPI000005256D; EMBL:Z50161; NID:G929846; PIDN:CAA90525.1; PI
C:Genetics:
A:Cross-references: SGD:S0005045
A:Map position: 14L
A:Note: YNL101w
C:Superfamily: Saccharomyces cerevisiae probable membrane protein YNL101w
C:Keywords: transmembrane protein
F:302-318/Domain: transmembrane #status predicted <TM1>
F:374-390/Domain: transmembrane #status predicted <TM2>
F:410-426/Domain: transmembrane #status predicted <TM3>
F:442-458/Domain: transmembrane #status predicted <TM4>

F:486-502/Domain: transmembrane #status predicted <TM5>
F:516-532/Domain: transmembrane #status predicted <TM6>
F:557-573/Domain: transmembrane #status predicted <TM7>
F:626-642/Domain: transmembrane #status predicted <TM8>
F:649-665/Domain: transmembrane #status predicted <TM9>
F:689-705/Domain: transmembrane #status predicted <TM10>

Query Match 18.8%; Score 465.5; DB 2; Length 713;

Best Local Similarity 31.0%; Pred. No. 1.3e-29;
Matches 144; Conservative 81; Mismatches 167; Indels 73; Gaps 18;

Qy 32 LSSPGSVQRQCSNST--WFQTLIHLKNGITGTLGLPLAVKNAGIVMGFISLLIIGIV 90

Db 281 LSRPDHMKVLPSAKGTSTTKKVFILKSPFIGTGVFLPNAFHNGGLFFSVSMIAFFGIY 340

Qy 91 AVHCWGLVLRCAHFCRRLNKSFDYDGTVMYGLSSPCSWLRNHAHGRVVDFFFLVIT 150

Db 341 SYWCYVILVQ--AKSSCG--VSSFGDIG-----LKLPGPMRIILFLSLVIT 383

Qy 151 QLGFCCVYFVLADNFQKQVIEAANGTTNNCHNETVILTPMDSRLYMLSFPLFLVL-LV 209

Db 384 QVSPGAYMIPTAKNLQAFLD-----NVFH-----VGLPLS-----YLMVFQTIIFILS 429

Qy 210 FIRNLRLSIFSLLANITMLSVLMIYQFIVQRI-PDPShLPVLA-----PWKTYPL 260

Db 430 FIRNISKLSPSLANFFIMAGLVIVIIFTAKRLFFDLMGTPAMGVVYGLNADRW---TL 486

Qy 261 PFGTAFSFGIGMVLPLENKKMDPKFPLIILYLGWVITVILYISLGLCYGLQFGANIQQ 320

Db 487 FIGTAIFAFEGILIIIPVQDSMRNPEKFLVLAIVLTATILFISIALGLAYGSNVQT 546

Qy 321 SITLNPCLWYQSVKLYSIGIFFTYALQFVVPABIIIPFVSVRAPEHCELV-- 373

Db 547 VILLNLQSNIFVNLIOLFYSIAIMLSPTLQLPAIKIENKFFPKFKIYVGHDDLTR 606

Qy 374 VDL-----FVRTVLVCLTCLAIL-IPRLDIVLSVCSVSSSALALIIIP 416

Db 607 VELRPNSGKLNWKIKLKNFIRSIIVVSIAYFGSDNLDKFKVSVIGSLACIPLVIYIP 666

Qy 417 PLLRVT--TFYSEGMSPLTIKDALISILGFVG---FVVGTYEALY 457

Db 667 SMLHRLGNSLPETKGEFWRFKPMLDITLIFFGIASMLYTSQSIF 711

RESULT 6

TI5799

hypothetical protein C44B7.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

C:Accession: TI5799

R:Du, Z.

submitted to the EMBL Data Library, June 1995

A:Description: The sequence of C. elegans cosmid C44B7.

A:Reference number: S61146

A:Accession: TI5799

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-434 <DUZ>

A:Cross-references: UNIPROT:Q18595; UNIPARC:UPI000017B834; EMBL:U28928; NID:G861301; PID

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:C44B7.6

A:Introns: 22/2; 52/3; 82/2; 113/2; 141/3; 196/3; 250/1; 293/2; 356/1

Query Match 18.8%; Score 464; DB 2; Length 434;

Best Local Similarity 30.1%; Pred. No. 1e-29;

Matches 125; Conservative 85; Mismatches 177; Indels 28; Gaps 9;

Qy 53 LIHLKNGIGLGLPLAVKNAGIVMGFISLLIIGIVAVHCVGMIVKCAHHFCRRLKS 112

Db 27 LINLMKMLGAGCSFVAFKQSGVSLVILVIGLFCALCMILKVCAGYLSKVNSA 86

Qy 113 FVDYGDVTVMYGLSSPCSWLRNHAHGRVVDFFLIVTQLGFCVCFVFLADNFKQVIEA 172

Db 87 PLYDGN-MAYKATQASYPTRKAPVSRALVNSLCLQLGICCCCFYFVYVHLHELLEF 145
QY 173 ANGTNNCHNETVILPTWDSRLYMLSPFLVLLVFIIRNLRALSIFSLANITMLVSL 232
Db 146 --VMDVPSRAT--LFPV-----LPAFILLVSLSSMRALSIVSLGSGFLMLIAL 191
QY 233 -VMYQFIVORI PDPHSLPVAIPWKTYPLPFGTAIFSFEIGVMVPLENKKDKPR--KFP 289
Db 192 AVIMFOLLTTEHKKLADLPVTDLMGIVSAAGTILVALEGOAMVPLENRMKPKEDMKGP 251
QY 290 L-ILYLGWVITLYISLGLCYLOFGANTQGSITLNPNCWLYQSVKLYSGIFPTYA 348
Db 252 FGVLSVGVGMVVIYSFAGFGLTYGNDVQDSITLNPNDHILGIFVKAVLLFVYVSGFL 311
QY 349 LQFVPAEIIIPFVSRAPEHC-----ELVVDLFVRTVLVCLTCLAILIPRLDIVISL 402
Db 312 IQVFPPIVAMIWPAIKKLRITCGVSITTKRIIVHFAFRYSIVVFLLSYAIPRLSDMVPL 371
QY 403 VGVSSSALALIIPPLLEVTFYSEGMSPLTIKDALISILGFVGVVGYTYEALY 457
Db 372 VGVTAGMLLALVPSLPHLLILPQPECRIGFLFDI---FLDFVCIILGMFFVIY 423
RESULT 7
T23131
hypothetical protein H32K16.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T23131
R;Wallis, J.
submitted to the EMBL Data Library, June 1997
A;Reference number: Z19691
A;Accession: T23131
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-481 <WIL>
A;Cross-references: UNIPROT:O62286; UNIPARC:UPI000007809E; EMBL:Z97191; PIDN:CAE10025.1;
A;Experimental source: clone H32K16
C;Genetics:
A;Gene: CESP:H32K16.1
A;Map position: 1
A;Introns: 64/2; 94/3; 124/2; 155/2; 183/3; 240/3; 337/2; 397/1; 444/2
C;Superfamily: Arabidopsis amino acid transport protein 1

Query Match 18.7%; Score 462.5; DB 2; Length 481;
Best Local Similarity 28.7%; Pred. No. 1.5e-29;
Matches 123; Conservative 91; Mismatches 174; Indels 41; Gaps 8;
QY 53 LIHLKGNIGTGLGLPLAVKNAGIVMGPISSLIIIGIVAVHGMILVKCAHFCRRLNKS 112
Db 69 LINFICWIGPGCFSLAVSFQAGLWGLALVFIVGFLSLYSMHKIVNCQYLAKSNGDQ 128
QY 113 FVDYGDVTWYGLSSPCSWLRNHAHWGRRVVDFFLIQTOLGFCVCYVFFVLADNPKQVIEA 172
Db 129 SLDYGEMAEAAQMS--YKWARHGKGLAKIVINACLLAFQLGVIITVFMVFAVEHVIWEF 187
QY 173 ANGTNNCHNETVILPTWDSRLYMLSPFLVLLVFIIRNLRALSIFSLANITMLVSL 232
Db 188 FADSP-----PPFSKCVMLMYFVP-QMLNFIIGHMKLLTILCLFGNVIIPAAI 235
QY 233 VMIYQFIVORIPDPS-HLPLVAPWKTYPLPFGTAIFSFEIGVMVPLENKKDKPRK- 288
Db 236 VLITKELMVHTWYPTWELGVSVTGIEGSLAAGALIYSFEGQAMVPLMENSILKYPKDMTGA 295
QY 289 PLIILYLGWVITLYISLGLCYLOFGANTQGSITLNPNCWLYQSVKLYSGIFPTYA 348
Db 296 TGVLSLTMNLVTLVLYAFLPFGFYVTFGPAVQGSITLNPNSILTVSIRKGLVLKIFGSA 355
QY 349 LQFVPAEIIIPFVSRAPEHCVELVDLF---VRTVLVCLTCLAILIPRLDIVISLVGS 405
Db 356 IQLYVIVQMLLPSLRKISDEKRWVHRLPLPVALRLGLMLSLCLIALIVPMLQIPLVGI 415

QY 406 VSSSALALIIPPLLEVTFYSEGMSPLTIK-----DALISILGFVGVFVWG 451
Db 416 TSGLLSILSLPSFLDCMVF-----LPVFKQOGDMFKFYQKLIINVLFLVGLWFFLGAG 468
QY 452 TYEALYELI 460
Db 469 LYSSIDII 477
RESULT 8
F88544
protein F59B2.2 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: F88544
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C_elegans/
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: F88544
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-460 <STO>
A;Cross-references: UNIPROT:P34479; UNIPARC:UPI000013B9DA; GB:chr_III; PIDN:CAA77582.1;
C;Genetics:
A;Gene: F59B2.2
A;Map position: 3

Query Match 18.0%; Score 443.5; DB 2; Length 460;
Best Local Similarity 26.7%; Pred. No. 4.7e-28;
Matches 128; Conservative 101; Mismatches 211; Indels 39; Gaps 10;
QY 6 LRNEDYHYSSTDVSPESPSEGLNMLSSPGSYORFGQSNSTTWFTOLIHLLKGNIGTGL 65
Db 1 MSEEGERARVTEGDAESMNDGRALVQPPA---RSGDVITPT--RAVLTLSKSMFNAGC 55
QY 66 LGPLAVKAGIYMGPISSLIIIGIVAVHGMILVKCAHFCRRLNKSFDYDGTVMYGLE 125
Db 56 FSLPYAWKGLGLVWSFVMSFVIAGLWNWYGNHILVRASQHLAKKSDRSALDYGHFAKKVCD 115
QY 126 SSPCSWLRNHAHWGRRVVDFFLIQT---OLGECVYFVFLADNPKQVIEAANGTTNNCH 181
Db 116 YSDIRFLRNS-----KAVMYFVNVTILFYOLGCSVAILFISDNLVNLVGDHLGGT--- 167
QY 182 NNETVILPTMDSRLYMLSPFLVLLVFIIRNLRALSIFSLANITMLVSLVMIYQFIVQ 241
Db 168 RHQOMILMATVS-----LFFILLTNMFTEMRIYVSFFALVSVFVIGAAVIMQYTVQ 219
QY 242 RIPDPHSLPVAIPWKTYPLPFGTAIFSFEIGVMVPLENKKDKPRK- -PL-ILYLGWVI 298
Db 220 QPNQWKLPAATNFTGTTMIGMSYAFEGQTMILPIENKLDNPAAFLAPFGLSTTMII 279
QY 299 VTILYISLGLCYLOFGANTQGSITLNPNCWLYQSVKLYSGIFPTYAQLQFVPAEII 358
Db 280 CTAFMTALGFFGTGFGDSIAPTITTNVPKEGLYSTVNVFLMQSLGNSIAMVYVVDMF 339
QY 359 I----PFVVSRAPEHCVELVDLFVRTVLVCLTCLAILIPRLDIVISLVGSVSSALALI 414
Db 340 FNGFRKRGARFNVFNKLSDKGRFVFWLVLMVAVLIPKLEIMIPLVGVTSGALCALI 399
QY 415 IPPLELVTFYSEGMSPLT-----IFKDALISILGFVGVVGYTYEALYELIQPSNAP 466
Db 400 FPPFEMITPWTDKGLLTYRQMTKIFINLVVMAIGVFAIAGVYTNHAI IQSFSP 458
RESULT 9
T26705
hypothetical protein Y38H6C.17 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T26705

A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-132 <VA3>
A;Cross-references: UNIPARC:UPI0000168BC6; EMBL:Z26877; NID:g407482; PIDN:CAA81508.1; PI
A;Experimental source: strain S288C
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1993
C;Genetics:
A;Cross-references: SGD:S0001629
A;Map position: 11L
A;Superfamily: Saccharomyces cerevisiae probable membrane protein YNL101W
C;Keywords: transmembrane protein

Query Match 15.7%; Score 388.5; DB 2; Length 692;
Best Local Similarity 28.9%; Pred. No. 1.9e-23;
Matches 131; Conservative 76; Mismatches 193; Indels 53; Gaps 14;

QY 8 NEDYDYSSTDVSPSPSGLNLSPPGYSQVQFGQSNSTWQTLLHLKGNIGTGLLG 67
Db SEEEEEETEEPEEALTESTQLVSRHGRPHKSTV--KAVLLLKSGFTGVLP 317
QY 68 LPLAVKNAGIVMGPIISLLIIGIVAVHCHMGLVKCAHFCRRLNKSFYD-YGD--TWMYGL 124
Db LPLAFHNGWGFSALCLLSCALISYGCFSVSLITK-----DKVGVDGYGDMGRILYGP 370
QY 125 ESSPSCSLRNHAWGRVRVDFLLIVTQGLCCVYFVPLADNFQVIEAANGTTNNCHNE 184
Db 371 KM-----KPAILLSALSQIGFSAAYTFTATNLQVFS-----NFFH--- 408
QY 185 TVILTP-TMDSRLYMLSFLLVLLVFIIRLALSIFSLANITMVLVSLVMYQFVQRI 243
Db 409 ---LKPGSISLTYFAQVLIIFVPLSLTRIAKLSGTALADUFIILGLVYVYIYI 455
QY 244 -----PDPSHLPLVAPWKTYPTFFGTAIFSGIGMVLPLENKKMDPKRFPILYLGMV 297
Db 466 AVNGVADTWMFNKADWS---LFIGTAIFTEGIGLLIPIQSMKHPKHFPSLSAVMC 522
QY 298 IVTLYTSLGCLGYLQGANIQGSITNLNPNWYQ-SVKLLYSIGIFTFYALQFYVPAAE 356
Db 523 IVAVIFISGCLLCVAAGSDVKTVLLENFPDTSYTLTVQLLYALAILLSTPLQLFPAIR 582
QY 357 IITPF-FVSRAPHECELVDL---FVRTVLVCLTCILA-LIPLDLVISLGVSSSAL 411
Db 583 ILENWTFFSNAGSKYNPKVKNYFCAIVLVLSILAWVGANDLDFVSLVGSFACIPL 642
QY 412 ALIIPPLEVTFYSEGMSPLTFKDALISTLG 444
Db 643 IYIYPLLLHKASILSGTSRARLLLDLIVIVFG 675

RESULT 12
S31123
hypothetical protein F59B2.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: S31123
R;Sulston, J.; Du, Z.; Thomas, K.; Willson, R.; Hillier, L.; Staden, R.; Halloran, N.; G
awkins, T.; Ainscough, R.; Waterston, R.
submitted to the EMBL Data Library, November 1991
A;Description: The C. elegans sequencing project: A beginning.
A;Reference number: S31122
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-389 <SUL>
A;Cross-references: UNIPROT:P34479; UNIPARC:UPI000017BA50; EMBL:Z11505; NID:g6718; PID:g
A;Introns: 38/2; 59/2; 148/2; 343/3

Query Match 14.9%; Score 368; DB 2; Length 389;
Best Local Similarity 28.5%; Pred. No. 4.7e-22;
Matches 97; Conservative 72; Mismatches 141; Indels 30; Gaps 7;

QY 145 FFLIVT---QLGFCVYFVLADNFQVIEAANGTTNNCHNETVILTPMTDSRLYMLS 200

Db 60 YFNVNITILFYQLGMCVAILFISDNLVNLVGDHLGCT----RHQMLLMATVS----- 108
QY 201 FLPLVLLVFIIRLALSIFSLANITMVLVSLVMYQFVQRIIPDPSPHLPVAPWKTYPL 260
Db 109 -LFFILLTNMFTMRIVSFVSFALVSSVFFVIGAIVMQYTVQPNQWDKLPAPATNFTGIT 167
QY 261 FPGTALFSEFEGIMVPLENKKMDPKF--PL-ILYLGWVIVTILYISLGLCYLQFAN 317
Db 168 MIGMSWYAFEGQTMILPIENKLDNPAFLAPFGVLSTTMIICITAFMTALGFGYTGFGDS 227
QY 318 IQGSITLNPNCWLYQSVKLLYSIGIFFTYALQFYVPAEIII---PFFVSRAPHECELV 373
Db 228 IAPTITNPNKGLYGVNVFLMLQSLGNSIAMVYVDMFFNGFRKFGARFNPVKWL 287
QY 374 VDLFVRTVLVCLTCILAAILPRDLVISLGVSSSALAILIPLLEVTTFYSEGMSPLT 433
Db 288 SDKGRFVFWVLYTLMVAVLIPKLEIMPLVGVTSGLCALIPFPFFEMITFTWDKGLLT 347
QY 434 -----IPKDALISILGFVGVVGYEALYELIQPSNAP 466
Db 348 YRQRMTKIFINLVMAIGVFAIAGVYTIHAIQSFQSP 387

RESULT 13
T48239
hypothetical protein T7H20.230 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C;Accession: T48239
R;Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; Mew
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24488
A;Accession: T48239
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-543 <BEV>
A;Cross-references: UNIPROT:Q9LZL4; UNIPARC:UPI00000A3EF5; EMBL:AL162508
A;Experimental source: cultivar Columbia; BAC clone T7H20
C;Genetics:
A;Map position: 5
A;Introns: 74/2; 168/1; 234/1; 242/1; 302/1; 375/2; 420/3; 476/1; 490/3; 511/3
A;Note: T7H20.230
C;Superfamily: Arabidopsis amino acid transport protein I

Query Match 11.6%; Score 285.5; DB 2; Length 543;
Best Local Similarity 25.3%; Pred. No. 2.9e-15;
Matches 123; Conservative 85; Mismatches 183; Indels 95; Gaps 21;

QY 2 STORLNEYDHYDYSST--DVSPEESPSEGLNNLSPPGYSQVQFGQSNSTWQTLLHLKLG 59
Db 119 SSKPILLSQPVDPKREETILPWNQSLKLSVTDLP-----EPNLCSFSQSVLNTV 171
QY 60 NITGTLGLPLAVKNAGIVMGPIISLLIIGIVAVHCHMGLVKCAHFCRRLNKS-----FV 114
Db 172 LCGLGLITMPYATIKESGWLGLPI-LLFVGIVTCY-TGLVLMK-----RLESPPGIQTP 223
QY 115 DYGDVTWYGLESPSCSLRNHAWGRVRVDFLLIVTQGLCCVYFVPLADN---FKQV- 169
Db 224 DIQAA-FGITDSI-----RGVVP-----CVEYIIMSDNLSGLFPNV 262
QY 170 IEAANGTTNNCHNETVILTPMTDSRLYMLSFLLVLLVFIIRLALSIFS---LLANI 226
Db 263 LSIASGIS-----LDSQPQIFAILTTLVLP---TWMLKDLSSLVSLVSGVGLASI 309
QY 227 TMLVSLVMYQ-----FIVQRIIPDPSPHLPVAPWKTYPLFFGTAFISFEGIGMVLPLEN 280
Db 310 LLGICLFWGAVDGI GFHATGRVFDLSNLPV-----IGIFGYSGHVSFPNIYS 360
QY 281 KMKDPKRPFLILYLGWVIVTILYISLGLCYLQFANIQGSITNLNPNWLYQSVKLLYS 340
Db 361 SMKDPSPFPLVLCISFCTVLIIVAVCCYTFMGEAVESQFTLNMKHPFPKVVAVMTA 420

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2006, 01:16:39 ; Search time 303 Seconds
(without alignments)
1453.160 Million cell updates/sec

Title: US-10-679-362-2
Perfect score: 2469
Sequence: 1 MSTQRLRNEDYHDSSTVDS.....YELIQPSNAPFINSTCAFI 476

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 92501592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 7.2.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2469	100.0	476	1 S36A1_HUMAN	Q722h8 homo sapien
2	2439	98.8	476	2 Q4R7M7_MACFA	Q47m7 macaca fasc
3	2317.5	93.9	475	2 Q2VPS4_RABIT	Q2vps4 oryctolagus
4	2146.5	86.9	475	1 S36A1_RAT	Q924a5 rattus norv
5	2141.5	86.7	475	1 S36A1_MOUSE	Q84id3 mus musculu
6	2141.5	86.7	475	2 Q5F227_MOUSE	Q5f227 mus musculu
7	1722.5	69.8	483	2 Q495M3_HUMAN	Q495m3 homo sapien
8	1713.5	69.4	483	2 Q6ZWK5_HUMAN	Q6zwk5 homo sapien
9	1711.5	69.3	483	2 Q726B5_HUMAN	Q726b5 homo sapien
10	1699.5	68.8	478	2 Q8BHK3_MOUSE	Q8bhk3 m adult mal
11	1696.5	68.7	478	2 Q8JZP1_MOUSE	Q8jzpl mus musculu
12	1689.5	68.4	479	2 Q6NRA6_XENLA	Q6nra6 xenopus lae
13	1677.5	67.9	481	2 Q8K415_RAT	Q8k415 rattus norv
14	1592	64.5	313	2 Q86YK4_HUMAN	Q86yk4 homo sapien
15	1476.5	59.8	470	2 Q726B4_HUMAN	Q726b4 homo sapien
16	1476.5	59.8	470	2 Q6ZRU4_HUMAN	Q6zru4 homo sapien
17	1472.5	59.6	470	2 Q495N2_HUMAN	Q495n2 homo sapien
18	1450.5	58.7	477	2 Q4V8B1_RAT	Q4v8b1 rattus norv
19	1447	58.6	511	2 Q495N3_HUMAN	Q495n3 homo sapien
20	1416.5	57.4	477	2 Q811P0_MOUSE	Q811p0 mus musculu
21	1415.5	57.3	477	2 Q8CH37_MOUSE	Q8ch37 mus musculu
22	1351	54.7	490	2 Q4S4A7_TETNG	Q4s4a7 tetraodon n
23	1301	52.7	455	2 Q6ZMU7_HUMAN	Q6zmu7 mus musculu
24	1275.5	51.7	500	2 Q8C077_MOUSE	Q8c077 mus musculu
25	1265.5	51.3	500	2 Q8CH36_MOUSE	Q8ch36 mus musculu
26	1255.5	50.9	504	2 Q6YBV0_HUMAN	Q6ybv0 homo sapien
27	1253	50.7	510	2 Q6DDP2_XENLA	Q6ddp2 xenopus lae
28	1253	50.7	522	2 Q4KL91_XENLA	Q4kl91 xenopus lae
29	1244.5	50.4	504	2 Q86X30_HUMAN	Q86x30 homo sapien
30	1080	43.7	285	2 Q495M4_HUMAN	Q495m4 homo sapien
31	1060.5	43.0	301	2 Q5R828_PONPY	Q5r828 pongo pygma

32	972.5	39.4	369	2 Q8N8S6_HUMAN	Q8n8s6 homo sapien
33	970.5	39.3	280	2 Q8BUB0_MOUSE	Q8bub0 mus musculu
34	866	35.1	482	2 Q8MSR2_DROME	Q8msr2 drosophila
35	866	35.1	504	2 Q9VLM3_DROME	Q9vlm3 drosophila
36	861	34.9	483	2 Q7KTI1_DROME	Q7kti1 drosophila
37	861	34.9	486	2 Q9VLM4_DROME	Q9vlm4 drosophila
38	853	34.5	426	2 Q7Q4M8_ANOGA	Q7q4m8 anopheles g
39	786.5	31.9	164	2 Q5F228_MOUSE	Q5f228 mus musculu
40	761	30.8	427	2 Q7QFG5_ANOGA	Q7qfg5 anopheles g
41	735	29.8	455	2 Q7PM43_ANOGA	Q7pm43 anopheles g
42	732	29.6	207	2 Q495M6_HUMAN	Q495m6 homo sapien
43	708	28.7	450	2 Q7Q199_ANOGA	Q7q199 anopheles g
44	703.5	28.5	486	2 Q8MU61_ACYPI	Q8mu61 acyrthosiph
45	701	28.4	468	2 Q9VX84_DROME	Q9vx84 drosophila

ALIGNMENTS

RESULT 1
S36A1_HUMAN STANDARD; PRT; 476 AA.
AC Q722H8; Q7Z7C0; Q96M74;
DT 16-AUG-2004, integrated into UniProtKB/Swiss-Prot.
DT 01-OCT-2003, sequence version 1.
DE Proton-coupled amino acid transporter 20.
DE DT 07-MAR-2006, entry version 20.
DE 1) (Solute carrier family 36 member 1).
GN Name=SLC36A1; Synonyms=PAT1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1).
RA Chen Z., Fei Y.-J., Huang W., Anderson C.M.H., Wake K.A.,
RA Thwaites D.T., Ganapathy V.;
RT "Structure and function of a proton-coupled amino acid transporter
(hPAT1) cloned from the human intestinal cell line Caco2";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), AND FUNCTION.
RC TISSUE=Intestine;
RX MEDLINE=22694810; PubMed=12809675; DOI=10.1016/S0888-7543(03)00099-5;
RA Boll M., Folz M., Rubio-Aliaga I., Daniel H.;
RT "A cluster of proton/amino acid transporter genes in the human and
mouse genomes";
RL Genomics 82:47-56(2003).
[3]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
RC TISSUE=Endometrial tumor;
RG The German cDNA consortium;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
[4]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 2).
RC PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Oobayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Matsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kinata M., Watanabe M., Hiraoaka S., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,

FT	CONFLICT	461	461	Q -> R (in Ref. 4).
SEQ	SEQUENCE	476 AA;	53076 MW;	628AE7FC7A6559P0 CRC64;
Query Match	100.0%;	Score 2469;	DB 1;	Length 476;
Best Local Similarity	100.0%;	Pred. No. 9.6e-17;		
Matches 476;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MSTQRLRNEDYHDYSDTSDYSPESPESPEGLNNLSSPGSYQRFQGSNSTTWFQTLIHLKGN	60	
Db	1	MSTQRLRNEDYHDYSDTSDYSPESPESPEGLNNLSSPGSYQRFQGSNSTTWFQTLIHLKGN	60	
Qy	61	IGTGLGLPLAVKNAGIVMGPIISLLIIGIVAVHGMGILVKCAHHFCKRLNKSFDVYGDV	120	
Db	61	IGTGLGLPLAVKNAGIVMGPIISLLIIGIVAVHGMGILVKCAHHFCKRLNKSFDVYGDV	120	
Qy	121	MYGLESPSCWLNHNAHGRVVVDFFELIVTQLGFCVCYFVFLADNFKQVIEAANGTTNNC	180	
Db	121	MYGLESPSCWLNHNAHGRVVVDFFELIVTQLGFCVCYFVFLADNFKQVIEAANGTTNNC	180	
Qy	181	HNNETVILPTMDSRLYMLSFPLFVLLVPIRNLRLALSTFLLANITMLVSLVMIYQFIV	240	
Db	181	HNNETVILPTMDSRLYMLSFPLFVLLVPIRNLRLALSTFLLANITMLVSLVMIYQFIV	240	
Qy	241	QRTPDPSHLPLVAPWKTYPLFCGTALFSPFEGIGMWLPLENKMDPKRFPPLILYLGWIVT	300	
Db	241	QRTPDPSHLPLVAPWKTYPLFCGTALFSPFEGIGMWLPLENKMDPKRFPPLILYLGWIVT	300	
Qy	301	ILYISLGLCYLQFGANIQGSITLNLPCNCLYQSVKLLYSIGIFTYALQFYVPAEIIIP	360	
Db	301	ILYISLGLCYLQFGANIQGSITLNLPCNCLYQSVKLLYSIGIFTYALQFYVPAEIIIP	360	
Qy	361	FFVSRAPHECELVVDLFFVRTLVLVCLICAILIPRLDLVLSVGSVSSSALALIIPPLE	420	
Db	361	FFVSRAPHECELVVDLFFVRTLVLVCLICAILIPRLDLVLSVGSVSSSALALIIPPLE	420	
Qy	421	VTTTFYSEGMSPLTIFKDALISILGFGVGVGTVEALYELIOPSNAPFINSTCAFI	476	
Db	421	VTTTFYSEGMSPLTIFKDALISILGFGVGVGTVEALYELIOPSNAPFINSTCAFI	476	
RESULT 2				
Q4R7M7	MACFA			
ID	Q4R7M7_MACFA	PRELIMINARY;	PRT;	476 AA.
AC	Q4R7M7;			
DT	19-JUL-2005,	integrated into UniProtKB/TrEMBL.		
DT	19-JUL-2005,	sequence version 1.		
DT	07-FEB-2006,	entry version 5.		
DE	Testis cDNA, clone: Q5A-14788,	similar to human solute carrier family		
DE	36 (proton/amino acid symporter), member 1 (SLC36A1)..			
OS	Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
OC	Cercopithecoidea; Cercopithecinae; Macaca.			
OX	NCBI_TaxId=9541;			
RN	[1]			
RN	NUCLEOTIDE SEQUENCE.			
RP	PubMed=15944441; DOI=10.1093/molbev/msi187;			
RX	Osada N., Hirata M., Tanuma R., Kusuda J., Hida M., Suzuki Y.,			
RA	Sugano S., Gojobori T., Shen C.-K.J., Wu C.I., Hashimoto K.;			
RT	"Substitution Rate and Structural Divergence of 5'UTR Evolution:			
RT	Comparative Analysis Between Human and Cynomolgus Monkey cDNAs.";			
RL	Mol. Biol. Evol. 22:1976-1982(2005).			
RL	[2]			
RP	NUCLEOTIDE SEQUENCE.			
RG	International consortium for macaque cDNA sequencing and analysis;			
RT	"DNA sequences of macaque genes expressed in brain or testis and its			
RT	evolutionary implications.";			
RL	Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.			
CC	Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms			
CC	Distributed under the Creative Commons Attribution-NoDerivs License			
CC	EWBL; AB168788; BAB00895.1; -; mRNA.			

DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005279; F:amino acid-polyamine transporter activity; IEA.
DR GO; GO:0006865; P:amino acid transport; IEA.
DR InterPro: IPR002422; AA/rel_permease2.
DR Pfam: PF01490; Aa_trans_1.
SQ SEQUENCE 476 AA; 53063 MW; 6485B3BBF94CREF8 CRC64;

Query Match 98.8%; Score 2439; DB 2; Length 476;
Best Local Similarity 98.3%; Pred. No. 1.4e-168;
Matches 468; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSTQRLNEDYHDYSTDVSPESSEGLNNLSSPGSYQRFQSNSTTFQTLIHLKGN 60
DB 1 MSTQRLNEDYHDYSTDVSPESSEGLNNLSSPGSYQRFQSNSTTFQTLIHLKGN 60

QY 61 IGTGLLGLPLAVKNAAGIVMGPTSLIIIGIVAVHCGILVKCAHFCRLNKSFDVYDGT 120
DB 61 IGTGLLGLPLAVKNAAGIVMGPTSLIIIGIVAVHCGILVKCAHFCRLNKSFDVYDGT 120

QY 121 MYGLESPCSWLNRNHAHGRVVDFFLIQTQGFCCVYFVFLADNFKQVIEAANGTTNCC 180
DB 121 MYGLESPCSWLNRNHAHGRVVDFFLIQTQGFCCVYFVFLADNFKQVIEAANGTTNCC 180

QY 181 HNNEVTILPTWDSRLYMLSPFLVLLVFIENRLALSIFSLANITMLVSLVMIYQFIV 240
DB 181 HNNEVTILPTWDSRLYMLSPFLVLLVFIENRLALSIFSLANITMLVSLVMIYQFIV 240

QY 241 QRIPDPSHLPLVAPWKTYPLFFGTAFIPEFEGIGMVLPLENKKDKPRKPLIILYLGWIVT 300
DB 241 QRIPDPSHLPLVAPWKTYPLFFGTAFIPEFEGIGMVLPLENKKDKPRKPLIILYLGWIVT 300

QY 301 ILYISLGLGYLQFGANIQGSITLNLPCNLQYQSVKLLYSIGIFFTYALQFVPAEIIIP 360
DB 301 ILYISLGLGYLQFGANIQGSITLNLPCNLQYQSVKLLYSIGIFFTYALQFVPAEIIIP 360

QY 361 FVSRAPHECELVDLFTVRLVCLTCLAILIPLDLVISLVGSVSSALALIIPPLE 420
DB 361 FVSRAPHECELVDLFTVRLVCLTCLAILIPLDLVISLVGSVSSALALIIPPLE 420

QY 421 VTFYSEGSPITIFKDALISILGFTGVGVGYEALYELIOPSNAPIFINSTCAFI 476
DB 421 VTFYSEGSPITIFKDALISILGFTGVGVGYEALYELIOPSNAPIFINSTCAFI 476

RESULT 3
Q2VPS4 RABIT PRELIMINARY; PRT; 475 AA..
AC -Q2VPS4;
DT 10-JAN-2006, integrated into UniProtKB/TrEMBL.
DT 10-JAN-2006, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Proton/amino acid transporter 1.
GN Name=PAT1;
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha; Leporidae;
OC Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Miyauchi S., Abbot E.L., Zhuang L., Subramanian R.K., Ganapathy V.,
RA Thwaites D.T.;
RT "Isolation and function of rabbit PAT1 (slc36a1) and coexpression with
RT the IMINO transporter in renal brush-border membrane vesicles.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ Databases.
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY989816; AAY42402.1; -; mRNA.
SQ SEQUENCE 475 AA; 52701 MW; A66BB738C7D2B27F CRC64;

Query Match 93.9%; Score 2317.5; DB 2; Length 475;

Best Local Similarity 92.2%; Pred. No. 9.3e-160;
Matches 439; Conservative 23; Mismatches 13; Indels 1; Gaps 1;

QY 1 MSTQRLNEDYHDYSTDVSPESSEGLNNLSSPGSYQRFQSNSTTFQTLIHLKGN 60
DB 1 MSTQRLNEDYHDYSTDVSPESSEGLNNLSSPGSYQRFQSNSTTFQTLIHLKGN 59

QY 61 IGTGLLGLPLAVKNAAGIVMGPTSLIIIGIVAVHCGILVKCAHFCRLNKSFDVYDGT 120
DB 61 IGTGLLGLPLAVKNAAGILMGPLSLIIVMGIVAVHCGILVKCAHFCRLNKSFDVYDGT 119

QY 121 MYGLESPCSWLNRNHAHGRVVDFFLIQTQGFCCVYFVFLADNFKQVIEAANGTTNCC 180
DB 120 MYGLESPSSWLNRNHAHGRVVDFFLIQTQGFCCVYFVFLADNFKQVIEAANGTTSDC 179

QY 181 HNNEVTILPTWDSRLYMLSPFLVLLVFIENRLALSIFSLANITMLVSLVMIYQFIV 240
DB 180 HNNEVTILPTWDSRLYMLSPFLVLLVFIENRLALSIFSLANITMLVSLVMIYQFIV 239

QY 241 QRIPDPSHLPLVAPWKTYPLFFGTAFIPEFEGIGMVLPLENKKDKPRKPLIILYLGWIVT 300
DB 240 QRIPNPSHLPLVAPWKTYPLFFGTAFIPEFEGIGMVLPLENKKDKPRKPLIILYLGWIVT 299

QY 301 ILYISLGLGYLQFGANIQGSITLNLPCNLQYQSVKLLYSIGIFFTYALQFVPAEIIIP 360
DB 300 ALYISLGLGYLQFGANIQGSITLNLPCNLQYQSVKLLYSIGIFFTYALQFVPAEIIIP 359

QY 361 FVSRAPHECELVDLFTVRLVCLTCLAILIPLDLVISLVGSVSSALALIIPPLE 420
DB 360 FVSRAPHECELVDLFTVRLVCLTCLAILIPLDLVISLVGSVSSALALIIPPLE 419

QY 421 VTFYSEGSPITIFKDALISILGFTGVGVGYEALYELIOPSNAPIFINSTCAFI 476
DB 420 ITTYSEGSPITIAKDALISILGFTGVGVGYEALYELIOPSNAPIFINSTCAV 475

RESULT 4
S36A1 RAT STANDARD; PRT; 475 AA..
AC Q92AA5;
DT 16-AUG-2004, integrated into UniProtKB/Swiss-Prot.
DT 01-DEC-2001, sequence version 1.
DT 07-MAR-2006, entry version 23.
DE Proton-coupled amino acid transporter 1 (Proton/amino acid transporter
DE 1) (Solute carrier family 36 member 1) (lysosomal amino acid
DE transporter 1) (LYAAT-1) (Neutral amino acid/proton symporter).
GN Name=Slc36a1; Synonyms=LYAAT1, Pat1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA], FUNCTION, SUBCELLULAR LOCATION, AND TISSUE
RP SPECIFICITY.
RC STRAIN=Sprague-Dawley; TISSUE=Hippocampus;
RX MEDLINE=21309917; PubMed=11390972; DOI=10.1073/pnas.121183499;
RA Sagne C., Agulhon C., Ravassard P., Darmon M., Hamon M.,
RA El Mestikawy S., Gashier B., Giros B.;
RA "Identification and characterization of a lysosomal transporter for
RT small neutral amino acids.";
RT Proc. Natl. Acad. Sci. U.S.A. 98:7206-7211(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA], FUNCTION, SUBCELLULAR LOCATION, AND TISSUE
RP SPECIFICITY.
RC TISSUE=Brain;
RX PubMed=12598615;
RA Wreden C.C., Johnson J., Tran C., Seal R.P., Copenhagen D.R.,
RA Reimer R.J., Edwards R.H.;
RT "The H-coupled electrogenic lysosomal amino acid transporter LYAAT1
RT localizes to the axon and plasma membrane of hippocampal neurons.";
RL J. Neurosci. 23:1265-1275(2003).
CC -!- FUNCTION: Neutral amino acid/proton symporter. Probably involved

CC in the efflux of lysosomal proteolysis products such as L-proline,
CC L-alanine and glycine from the cell. May play a role in specifying

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Lysosomal and
CC plasma membrane. In neurons, colocalizes with the exocyst complex
CC in the axonal processes.

CC -1- TISSUE SPECIFICITY: Widely expressed and predominantly expressed
CC in brain. Within the brain, expression restricted to neurons and
CC not detected in glial cells. Abundant in regions rich in neurons
CC using glutamate and GABA such as Purkinje cells in the cerebellum
CC and pyramidal cells in the hippocampus.

CC -1- SIMILARITY: Belongs to the amino acid/polyamine transporter 2
CC family.

CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License

CC -----

DR EMBL; AF361239; AAK67316.1; -; mRNA.

DR Ensembl; ENSRNOG00000123356; Rattus norvegicus.

DR RGD; 619801; Slc36a1.

DR -GO; GO:0005764; C:lysosome; IDA.

DR -GO; GO:0005886; C:plasma membrane; IDA.

DR -GO; GO:0015078; F:hydrogen ion transporter activity; IDA.

DR -GO; GO:0015175; F:neutral amino acid transporter activity; IDA.

DR -GO; GO:0015804; P:neutral amino acid transport; IDA.

DR -GO; GO:0015992; P:proton transport; IDA.

DR InterPro; IPR002422; AA/rel_permease2.

DR Pfam; PF01490; Aa_trans_1.

KW Amino-acid transport; Glycoprotein; Membrane; Symport; Transmembrane;
KW Transport.

FT CHAIN 1 475 Proton-coupled amino acid transporter 1.
FT /FTID=PRO 0000093827.

FT TRANSMEM 51 71 Potential.

FT TRANSMEM 78 98 Potential.

FT TRANSMEM 141 161 Potential.

FT TRANSMEM 190 210 Potential.

FT TRANSMEM 215 235 Potential.

FT TRANSMEM 257 277 Potential.

FT TRANSMEM 289 309 Potential.

FT TRANSMEM 342 362 Potential.

FT TRANSMEM 372 392 Potential.

FT TRANSMEM 397 417 Potential.

FT TRANSMEM 439 459 Potential.

FT CARBOHYD 43 43 N-linked (GlcNAc...) (Potential).

FT CARBOHYD 133 133 N-linked (GlcNAc...) (Potential).

FT CARBOHYD 173 173 N-linked (GlcNAc...) (Potential).

FT CARBOHYD 182 182 N-linked (GlcNAc...) (Potential).

FT CARBOHYD 469 469 N-linked (GlcNAc...) (Potential).

FT CONFLICT 108 108 L -> M (in Ref. 2).

SQ SEQUENCE 475 AA; 52569 MW; D3D04A489AD1D23C CRC64;

Query Match 86.9%; Score 2146.5; DB 1; Length 475;

Best Local Similarity 85.7%; Pred. No. 2.4e-147;

Matches 408; Conservative 36; Mismatches 31; Indels 1; Gaps 1;

QY 1 MSTQRLNEDVHDYSSTDVSPESRGLNLLSPGYSQRFQGSNSTWFTQTLHLKGN 60

DB 1 MSTQRLNEDVHDYSSTDVSPESRGLGSPGYSQRLGNSMTWFTQTLHLKGN 59

QY 61 IGTGLLGLPLAVKNAGIVMGPIISLLIIGIVAVHGMGILVKCAHHFCRLNKSFDYDGTV 120

DB 60 IGTGLLGLPLAVKNAGLLGLPLSLVIGIVAVHGMGILVKCAHHLCLRLNKPFLDYDGTV 119

QY 121 MYGLESPCSWLNRHIAHWGRVDFLLIVTQLGCCVYFVLADNFKQVTEAANGTTNC 180

DB 120 MYGLESPSTWIRNHSWGRVDFLLVWTQLGCCVYFVLADNFKQVTEAANGTTTC 179

QY 181 HNNETVILTPMDSRLYMLSLPFLVLLVFLTRNLRLSIFSLANIIMVSLVMIYQFIV 240

DB 180 NNNETVILTPMDSRLYMLTFLPLVLLSFLTRNLRLSIFSLANISMFVSLMIYQFIV 239

QY 241 QRIPDPRLPLVAPWKTYPLFFGTATFISFEGIGMVLPLENKKMDPRKPLILYLGMVIT 300

Db 240 QRIPDPRLPLVAPWKTYPLFFGTATFISFEGIGMVLPLENKKMDSKQKFLILYLGMVIT 299

QY 301 ILYISGCGLYQFGANIQGSITLNPNCWLQSVKLYSIGIFFYALQFYVPAEIIIP 360

Db 300 VLYISLGLYLQFGADIKGSITLNPNCWLQSVKLYSIGIFFYALQFYVPAEIIIP 359

QY 361 FVFSRAPEHCELVVDLFTVTLVCLTCILAILPRLDLVLISLVGSSSSALALIIPLE 420

Db 360 AIVSRVPERFELVDLSARTAMCVTCVLAVLPRLDLVLISLVGSSSSALALIIPLE 419

QY 421 VTFYSEGSPITIPKDALISILGFVGVVGTVEALYELIQSPNAPIFINSTCAFI 476

Db 420 VTFYVEGSPITITKDALISILGFVGVVGTVEALYELIQSPSHSDSTNSTSAFI 475

RESULT 5

S36A1_MOUSE

ID S36A1_MOUSE STANDARD; PRT; 475 AA.

AC Q8K4D3; Q811N9;

DT 16-AUG-2004, integrated into UniProtKB/Swiss-Prot.

DT 01-OCT-2002, sequence version 1.

DT 07-MAR-2006, entry version 21.

DE Proton-coupled amino acid transporter 1 (Proton/amino acid transporter

DE 1) (Solute carrier family 36 member 1).

GN Name=Slc36a1; Synonyms=Pat1;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP NUCLEOTIDE SEQUENCE [MRNA], FUNCTION, AND TISSUE SPECIFICITY.

RC STRAIN=C57BL/6;

RX MEDLINE=22063354; PubMed=11959859; DOI=10.1074/jbc.M200374200;

RA Boll M., Foltz M., Rubio-Allaga I., Kottira G., Daniel H.;

RT "Functional characterization of two novel mammalian electrogenic

RT proton-dependent amino acid cotransporters.";

RL J. Biol. Chem. 277:22966-22973(2002).

RN [2]

RP NUCLEOTIDE SEQUENCE [MRNA].

RX PubMed=15058382; DOI=10.1007/s00335-003-2319-3;

RA Birmingham J.R. Jr., Pennington J.;

RT "Organization and expression of the SLC36 cluster of amino acid

RT transporter genes.";

RL Mamm. Genome 15:114-125(2004).

CC -1- FUNCTION: Neutral amino acid/proton symporter. Has a pH-dependent

CC electrogenic transport activity for small amino acids such as

CC glycine, alanine and proline. Besides small apolar L-amino acids,

CC it also recognize their D-enantiomers and selected amino acid

CC derivatives such as gamma-aminobutyric acid.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Lysosomal and

CC plasma membrane (By similarity).

CC -1- TISSUE SPECIFICITY: Highly expressed in small intestine, colon,

CC kidney and brain.

CC -1- SIMILARITY: Belongs to the amino acid/polyamine transporter 2

CC family.

CC -----

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NoDerivs License

CC -----

DR EMBL; AF453743; AAM80480.1; -; mRNA.

DR EMBL; AY211262; AAC37090.1; -; mRNA.

DR Ensembl; ENSMUSG0000020261; Mus musculus.

DR MGI; MGI:2445299; Slc36a1

DR GO; GO:0015187; F:glycine transporter activity; IDA.

DR GO; GO:0015180; F:L-alanine transporter activity; IDA.

DR GO; GO:0015193; F:L-proline transporter activity; IDA.

DR GO; GO:0015816; P:glycine transport; IDA.

DR GO; GO:0015808; P:L-alanine transport; IDA.

DR GO; GO:0015824; P:L-proline transport; IDA.

DR GO; GO:0015992; P:proton transport; IDA.

DR InterPro; IPR02422; AA/rel_permease2.


```
DR InterPro; IPR013057; AA transp_TM.
KW Pfam; PF01490; AA trans; 1.
KW Amino-acid transport; Glycoprotein; Membrane; Sympor; Transmembrane;
KW Transport.
FT CHAIN 1 475 Proton-coupled amino acid transporter 1.
FT /FTID=PRO_0000093826.
FT TRANSMEM 51 71 Potential.
FT TRANSMEM 78 98 Potential.
FT TRANSMEM 141 161 Potential.
FT TRANSMEM 190 210 Potential.
FT TRANSMEM 215 235 Potential.
FT TRANSMEM 257 277 Potential.
FT TRANSMEM 289 309 Potential.
FT TRANSMEM 342 362 Potential.
FT TRANSMEM 372 392 Potential.
FT TRANSMEM 397 417 Potential.
FT TRANSMEM 439 459 Potential.
FT CARBOHYD 43 43 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 133 133 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 173 173 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 182 182 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 469 469 N-linked (GlcNAc...) (Potential).
FT CONFLICT 43 43 N -> S (in Ref. 2).
FT CONFLICT 198 198 L -> P (in Ref. 2).
SQ SEQUENCE 475 AA; 52466 MW; 113C23309F2B51F1 CRC64;

Query Match 86.7%; Score 2141.5; DB 1; Length 475;
Best Local Similarity 85.9%; Pred. No. 5.5e-147;
Matches 409; Conservative 33; Mismatches 33; Indels 1; Gaps 1;

QY 1 MSTQRLRNEDYHDSSTDVSPSEPSSEGLNNLSFGSYQRFQSGNSTTWFQTLIHLKGN 60
DB 1 MSTQRLRNEDYHDSSTDVSPSEPSSEGLGSPF-SPGSYQRLGENSMTWTFQTLIHLKGN 59

QY 61 IGTGLLGLPLAVKAGIWMGPIISLLIIGIVAVHGMGILVKCAHHFCRLNKSFDVYGDIV 120
DB 60 IGTGLLGLPLAVKAGLGLPLSLVIGIVAVHGMGILVKCAHHLCRLNKPFLDYGDIV 119

QY 121 MYGLESPCSWLNRHAWGRVDFPLVLTQLGCCVYFVLADNFKQVIEAANGTTNCC 180
DB 120 MYGLECSPTWVRNHSWGRIVDFLLVLTQLGCCVYFVLADNFKQVIEAANGTTTNC 179

QY 181 HNNETVILPTMDSRLYMLSELPFLVLVLFIRNLRALSIFSLANITMLVSLVMIYQFIV 240
DB 180 NNNVTVIPTMDSRLYMLSELPFLVLVLSFIRNLRVLSIFSLANISMVFSLSMIYQFIV 239

QY 241 QRIPDPSHLPLVAPWKTYPLFFGTAFPSFEGIGMVLPLENKKDKPRKPLIYLGMVIT 300
DB 240 QRIPDPSHLPLVAPWKTYPLFFGTAFPSFEGIGVVLPLENKKDKSQKPELIYLGMAIT 299

QY 301 ILVYSLGCLGYLQFAGNIQSGITLNLPCWLYQSVKLLYSIGIFPTYALQFVPAEIIIP 360
DB 300 VLYISLGLGYLQFAGNIKGSITLNLPCWLYQSVKLLYSIGIFPTYALQFVPAEIIIP 359

QY 361 FFVSRAPHCHELVVDLFFVTVLVCITLAILIPRLDLVISLVGSSSSALALIIPPLLE 420
DB 360 AIVSRVPEHFMWDLVLRVTAMVCTVLAIIIPRLDLVISLVGSSSSALALIIPPLLE 419

QY 421 VTTYFSEGMSPLTIKDALISILGFVGVGTGYEALYELIQSPNAPIFINSTCAPI 476
DB 420 VVTYVYEGISPLTVTKDALISILGFVGVGTGYESLCELIQPSHSDSSTNSTSAFI 475

RESULT 6
Q5F227 MOUSE
ID Q5F227 MOUSE PRELIMINARY; PRT; 475 AA.
AC Q5F227.
DT 10-MAY-2005, integrated into UniProtKB/TrEMBL.
DE 07-FEB-2006, sequence version 1.
GN Solute carrier family 36 (proton/amino acid symporter), member 1.
OS Mus musculus (Mouse).
```

```
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11
RP NUCLEOTIDE SEQUENCE.
RA Phillimore B.;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AL713870; CA152048.1; -; Genomic DNA.
DR GO; ENSMUSG00000020261; Mus musculus.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005279; P:amino acid-polyamine transporter activity; IEA.
DR GO; GO:0006865; P:amino acid transport; IEA.
DR InterPro; IPR002422; AA/rel_permease2.
DR InterPro; IPR013057; AA_transp_TM.
DR Pfam; PF01490; AA_trans; 1.
SQ SEQUENCE 475 AA; 52466 MW; 113C23309F2B51F1 CRC64;

Query Match 86.7%; Score 2141.5; DB 2; Length 475;
Best Local Similarity 85.9%; Pred. No. 5.5e-147;
Matches 409; Conservative 33; Mismatches 33; Indels 1; Gaps 1;

QY 1 MSTQRLRNEDYHDSSTDVSPSEPSSEGLNNLSFGSYQRFQSGNSTTWFQTLIHLKGN 60
DB 1 MSTQRLRNEDYHDSSTDVSPSEPSSEGLGSPF-SPGSYQRLGENSMTWTFQTLIHLKGN 59

QY 61 IGTGLLGLPLAVKAGIWMGPIISLLIIGIVAVHGMGILVKCAHHFCRLNKSFDVYGDIV 120
DB 60 IGTGLLGLPLAVKAGLGLPLSLVIGIVAVHGMGILVKCAHHLCRLNKPFLDYGDIV 119

QY 121 MYGLESPCSWLNRHAWGRVDFPLVLTQLGCCVYFVLADNFKQVIEAANGTTNCC 180
DB 120 MYGLECSPTWVRNHSWGRIVDFLLVLTQLGCCVYFVLADNFKQVIEAANGTTTNC 179

QY 181 HNNETVILPTMDSRLYMLSELPFLVLVLFIRNLRALSIFSLANITMLVSLVMIYQFIV 240
DB 180 NNNVTVIPTMDSRLYMLSELPFLVLVLSFIRNLRVLSIFSLANISMVFSLSMIYQFIV 239

QY 241 QRIPDPSHLPLVAPWKTYPLFFGTAFPSFEGIGMVLPLENKKDKPRKPLIYLGMVIT 300
DB 240 QRIPDPSHLPLVAPWKTYPLFFGTAFPSFEGIGVVLPLENKKDKSQKPELIYLGMAIT 299

QY 301 ILVYSLGCLGYLQFAGNIQSGITLNLPCWLYQSVKLLYSIGIFPTYALQFVPAEIIIP 360
DB 300 VLYISLGLGYLQFAGNIKGSITLNLPCWLYQSVKLLYSIGIFPTYALQFVPAEIIIP 359

QY 361 FFVSRAPHCHELVVDLFFVTVLVCITLAILIPRLDLVISLVGSSSSALALIIPPLLE 420
DB 360 AIVSRVPEHFMWDLVLRVTAMVCTVLAIIIPRLDLVISLVGSSSSALALIIPPLLE 419

QY 421 VTTYFSEGMSPLTIKDALISILGFVGVGTGYEALYELIQSPNAPIFINSTCAPI 476
DB 420 VVTYVYEGISPLTVTKDALISILGFVGVGTGYESLCELIQPSHSDSSTNSTSAFI 475

RESULT 7
Q495M3 HUMAN
ID Q495M3 HUMAN PRELIMINARY; PRT; 483 AA.
AC Q495M3.
DT 13-SEP-2005, integrated into UniProtKB/TrEMBL.
DE 07-FEB-2006, entry version 2.
GN Solute carrier family 36 (proton/amino acid symporter), member 2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
```

```

RN RP NUCLEOTIDE SEQUENCE.
RC TISSUE=PCR rescued clones;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,
RA Diatchenko L., Marusina K., Rubin A.A., Rubin G.M., Hong L.,
RA Stapleton M.J., Uedin T.B., Tohiyuki S., Carninci P., Prange C.,
RA Brownstein M.J., McKernan K.J., Malek J.A., Gumaratne P.H.,
RA Bosak S.A., Morley P.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Richards S., Wozney K.C., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
RN RP NUCLEOTIDE SEQUENCE.
RC TISSUE=PCR rescued clones;
RG NIH MGC Project;
RL Submitted (AUG-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
DR EMBL: BC101103; AAI01102.1; -; mRNA.
DR EMBL: BC101101; AAI01102.1; -; mRNA.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005279; F:amino acid-polyamine transporter activity; IEA.
DR GO: GO:0006865; P:amino acid transport; IEA.
DR InterPro: IPR002422; AA/rel_permease2.
DR Pfam: PF01490; Aa trans; 1.
SQ SEQUENCE 483 AA; 53216 MW; E283B8C6FC0C63666 CRC64;

Query Match 69.8%; Score 1722.5; DB 2; Length 483;
Best Local Similarity 71.9%; Pred. No. 1.4e-116;
Matches 340; Conservative 50; Mismatches 62; Indels 21; Gaps 4;

QY 2 STORLNRDHYDYSSTDVSPSESPSE--GLNNLSSPGSYQRFQGSNSTTWFQTLIHLKLG 59
DB 26 SAKLENKD-----STFL--DESPSESAGLK-----KTGKITVFQALHLVKG 66
QY 60 NIGTGLGLPLAVKNAGIVMGPISELLIIGIVAVHVMGILVKCAHFCRRLNKSPVDYGD 119
DB 67 NMGTGILGLPLAVKNAGILMGPLSLVWVGFIACHMHLVKCAQRFCKRLNKPMDYGD 126
QY 120 VMYGLESSPCSWLRNHAHMGRRVDFLIIVTQGFCCVYFVLADNPKQVIEAANGTNN 179
DB 127 VMHGLEANPAWLQNHAGHRHIVSFLLIITQLGFCVYIVFLADNLKQVVEAVNSTNN 186
QY 180 CHNNETVILPTMDSRLYMSFLPFLVLLVIRNLRLAISFLSLANITMLVSLVMIYQFI 239
DB 187 CYSNETVILPTMDSRLYMSFLPFLVLLVIRNLRLITFISMLANISMLVSLVLIQYI 246
QY 240 VQRIPOPSHLPLVAPWKTYPLFFGTAIFSPFEGIGMVLPLENKMCKPRKFFLILYGMVIV 299
DB 247 TQBIPOPSRPLVASWKTYPLFFGTAIFSPFEGIGMVLPLENKMCKNARHFFAILSLGMSIV 306
QY 300 TILYISLGCGLGYLOFGANIOGSITLNPNCWLYQSVKLLYSIGIFFYALQFYVPAEIII 359
DB 307 TSLYIGMAALGYRFGDDIKASISLNPNCWLYQSVKLLYIAGLCLCYALQFYVPAEIII 366
QY 360 PFFVSRAPHELCVVDLFRVTLVCLTICILAILPRDLVLISLVGSSSALAILIIPLL 419
DB 367 PFAISRSTRWALPDLISIRLVWVCLTCLLAILPRDLVLISLVGSSVGTALAILIIPLL 426

```

```

QY 420 EVTTFYSEGMSPLTIKDALISILGFGVFGVGYEALYELIQSNAPIFINST 472
DB 427 EVTTFYSEGMSPLTIKDALISILGFGVFGVGYEALYELIQSNAPIFINST 479

RESULT 8
Q6ZWK5_HUMAN PRELIMINARY; PRT; 483 AA.
AC Q6ZWK5;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 21-FEB-2006, entry version 13.
DE CDNA FLJ16051 fis, clone KIDNE2000832.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RA Tanigami A., Fujiwara T., Shibahara T., Goto Y., Hirao M., Shimizu F.,
RA Wakebe H., Ono T., Hishigaki H., Watanabe T., Ozaki K., Sugiyama T.,
RA Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J.,
RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,
RA Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K.,
RA Masuho Y., Nagai K., Isoqai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
DR EMBL: AK122630; BAC85496.1; -; mRNA.
DR Ensembl: ENSG00000186335; Homo sapiens.
DR HGNC: HGNC:18762; SLC36A2.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005279; F:amino acid-polyamine transporter activity; IEA.
DR GO: GO:0006865; P:amino acid transport; IEA.
DR InterPro: IPR002422; AA/rel_permease2.
DR InterPro: IPR013057; AA_transpt_1M.
DR Pfam: PF01490; Aa trans; 1.
SQ SEQUENCE 483 AA; 53244 MW; B29A1671ED6C867D CRC64;

Query Match 69.4%; Score 1713.5; DB 2; Length 483;
Best Local Similarity 71.5%; Pred. No. 6.3e-116;
Matches 338; Conservative 51; Mismatches 63; Indels 21; Gaps 4;

QY 2 STORLNRDHYDYSSTDVSPSESPSE--GLNNLSSPGSYQRFQGSNSTTWFQTLIHLKLG 59
DB 26 SAKLENKD-----STFL--DESPSESAGLK-----KTGKITVFQALHLVKG 66
QY 60 NIGTGLGLPLAVKNAGIVMGPISELLIIGIVAVHVMGILVKCAHFCRRLNKSPVDYGD 119
DB 67 NMGTGILGLPLAVKNAGILMGPLSLVWVGFIACHMHLVKCAQRFCKRLNKPMDYGD 126
QY 120 VMYGLESSPCSWLRNHAHMGRRVDFLIIVTQGFCCVYFVLADNPKQVIEAANGTNN 179
DB 127 VMHGLEANPAWLQNHAGHRHIVSFLLIITQLGFCVYIVFLADNLKQVVEAVNSTNN 186
QY 180 CHNNETVILPTMDSRLYMSFLPFLVLLVIRNLRLAISFLSLANITMLVSLVMIYQFI 239
DB 187 CYSNETVILPTMDSRLYMSFLPFLVLLVIRNLRLITFISMLANISMLVSLVLIQYI 246
QY 240 VQRIPOPSHLPLVAPWKTYPLFFGTAIFSPFEGIGMVLPLENKMCKPRKFFLILYGMVIV 299
DB 247 TQBIPOPSRPLVASWKTYPLFFGTAIFSPFEGIGMVLPLENKMCKNARHFFAILSLGMSIV 306
QY 300 TILYISLGCGLGYLOFGANIOGSITLNPNCWLYQSVKLLYSIGIFFYALQFYVPAEIII 359

```

Db	307	TSLYIGMAALGLYRFGDDIKASISLNPCNMDYQSVKLYTAGICTTALQFYVPAEII	366
Qy	360	PFFVSRAPHELCENVVDLFRVTLVCLTCLAILPRDLVLSLGVSSVSSALALIIPLLT	419
Db	367	PFAISRVSRWALPDLSIRLVVMVCLTCLLAILPRDLVPLVGSVSGTALALIIPLLT	426
Qy	420	EVTTYSEGMSPLTFKDALISILGFGVVGTYEALYELIQPSNAPINST	472
Db	427	EVTTYSEGMSPLTFKDALISILGFGVVGTYOALDELLKSESHUPSPNST	479

RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugiura K., Sultana R., Takenaka Y., Taki K.,
RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., Van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA Grimmer S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
RA Wahlstedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA Kashiwagi T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
RA Nishio T., Okada M., Plessey C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome.";
RL Science 309:1559-1563(2005).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J, and NOD; TISSUE=Activated spleen, Aorta and vein,
RC Dienecephalon, and Whole body;
RX PubMed=16141073; DOI=10.1126/science.1112009;
RG _RIKEN Genome Exploration Research Group, and Genome Science Group
RG (Genome Network Core Team) and the FANTOM Consortium;
RT "Antisense Transcription in the Mammalian Transcriptome.";
RL Science 309:1564-1566(2005).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J, and NOD; TISSUE=Activated spleen, Aorta and vein,
RC Dienecephalon, and Whole body;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato R., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gasterland T., Gariboldi M., Glasi C., Godzik A., Gough J.,
RA Grimmer S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagasima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Varaldo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J, and NOD; TISSUE=Activated spleen, Aorta and vein,
RC Dienecephalon, and Whole body;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyoko-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J, and NOD; TISSUE=Activated spleen, Aorta and vein,
RC Dienecephalon, and Whole body;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RL prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [7]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J, and NOD; TISSUE=Activated spleen, Aorta and vein,
RC Dienecephalon, and Whole body;
RX MEDLINE=20330913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kishinai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Imoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-Format
RL sequencing pipeline with 384 multicapillary sequencer";
RL Genome Res. 10:1757-1771(2000).
RN [8]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [9]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FVB/N; TISSUE=Salivary gland;
RX MEDLINE=22288257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., Kerneran K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

Query Match 68.8%; Score 1699.5; DB 2; Length 478;
Best Local Similarity 72.6%; Pred. No. 6.4e-115;
Matches 329; Conservative 50; Mismatches 67; Indels 7; Gaps 1;

```
Qy 20 SPEESPSEGLNLSPPGYSQRFQGSNSTTFTQTLHLKGNIGTGLGLPLAVKNAGIVM 79
Db 29 SQDSPANGSSSESSK-----KTGKITGQTLVHLVKGNGTGILGLPLAVKNAGILM 81
Qy 80 GPISLLIIGIVAVHGMILVKCAHFCRLNKSFDYGDVTVMYGLSPCSWLNRNHAHWG 139
Db 82 GPLSLLVWGLIACHMHLVRCAQRFCHRLNKPMDYGDVTVMHGLAFSPNWLQNHAWG 141
Qy 140 RRVVDFFLIVTQGFCCVYFVFLADNPKQVTEAANGTNNCHNETVLTPTMDSRLYML 199
Db 142 RRVVSFFLIVTQGFCCVYFVFLADNLKQVVEAVNSTISCHKNETVLTPTMDSRLYML 201
Qy 200 SFLPFLVLLVFNRLRALSIFSLANITMVLVSLVMYQFIVQRIIDPSPHLPVAPWKTY 259
Db 202 SFLPVLGLLVFNRLRALTIFSLANISMLVSLVIAQYIIQEIIPDASQLPLVASWKTY 261
Qy 260 LFFGTAFSPGIGMVLPLENKMCKPRKFPILLYLGMVITVILYISLGLYLOFGANIQ 319
Db 262 LFFGTAFSPGIGMVLPLENKMCKDARFPITLSLGMSTITLYIAIGALGYLRFDDIK 321
Qy 320 GSITLNLPCNWLQSVKLLYIGIFFTYALQFYVPAEIIIPFVSRAPEHCELVDLFR 379
Db 322 ASITLNLPCNWLQSVKLLYVVGILCTYALQFYVPAEIIIPLASQVSKRWALPVDLSIR 381
Qy 380 TVLVCLTCILAILPRDLVLSVGSVSSSALALIIPLEVTTFYSGMSPLTIFKDAL 439
Db 382 LALVCLTCMLAILPRDLVLSVGSVSSSALALIIPLELVTVYVYGGISPLATVKDAL 441
Qy 440 ISILGFGVGVVGTVEALYELIQTSPNAPFINST 472
Db 442 ISILGFGVGVVGTVEALDELKSGNSPALSNST 474

RESULT 11
Q8J2P1_MOUSE
ID Q8J2P1_MOUSE PRELIMINARY; PRT; 478 AA.
AC Q8J2P1;
DT 01-OCT-2002, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2002, sequence version 1.
DT 07-FEB-2006, entry version 16.
DE Trandorin 1 (proton/amino acid transporter 2).
GN Name=Sic36a2; Synonyms=PAT2, Trandl;
OS -Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN -[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Pou3f1 mutant sciatic nerves;
RX MEDLINE=223339849; PubMed=12451123;
RA Bermingham J.R. Jr., Shumas S., Whisenhunt T., Sirkowski E.E.,
RA O'Connell S., Scherer S.S., Rosenfeld M.G.;
RT "Identification of genes that are downregulated in the absence of the
RT POU domain transcription factor pou3f1 (Oct-6, Tst-1, SCIP) in sciatic
RT nerve.";
RL J. Neurosci. 22:10217-10231(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6;
RX MEDLINE=22063354; PubMed=11959859; DOI=10.1074/jbc.M200374200;
RA Boll M., Foltz M., Rubio-Aliaga I., Kottar G., Daniel H.;
RT "Functional characterization of two novel mammalian electrogenic
RT proton-dependent amino acid cotransporters.";
RL J. Biol. Chem. 277:22966-22973(2002).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF512429; AA04854.1; -; mRNA.
DR EMBL; AF453744; AA060481.1; -; mRNA.
DR Ensembl; ENSMUSG0000020264; Mus musculus.
DR MGI; MGI:1891430; Sic36a2.
```

```
DR GO; GO:0016021; C:integral to membrane; RCA.
DR GO; GO:0005886; C:plasma membrane; IDA.
DR GO; GO:0005279; F:amino acid-polyamine transporter activity; RCA.
DR GO; GO:0015187; F:glycine transporter activity; IDA.
DR GO; GO:0005280; F:hydrogen:amino acid symporter activity; IDA.
DR GO; GO:0015180; F:L-alanine transporter activity; IDA.
DR GO; GO:0015193; F:L-proline transporter activity; IDA.
DR GO; GO:0006865; P:amino acid transport; RCA.
DR GO; GO:0015816; P:glycine transport; IDA.
DR GO; GO:0015808; P:L-alanine transport; IDA.
DR GO; GO:0015824; P:L-proline transport; IDA.
DR GO; GO:0015932; P:proton transport; IDA.
DR InterPro; IPR002422; AA/rei_permease2.
DR InterPro; IPR013057; AA transport_TM.
DR Pfam; PF01490; Aa_trans; 1.
SQ SEQUENCE 478 AA; 52049 MW; 2EC02C59577BE236 CRC64;

Query Match 68.7%; Score 1696.5; DB 2; Length 478;
Best Local Similarity 72.6%; Pred. No. 1.1e-114;
Matches 329; Conservative 49; Mismatches 68; Indels 7; Gaps 1;

Qy 20 SPEESPSEGLNLSPPGYSQRFQGSNSTTFTQTLHLKGNIGTGLGLPLAVKNAGIVM 79
Db 29 SQDSPANGSSSESSK-----KTGKITGQTLVHLVKGNGTGILGLPLAVKNAGILM 81
Qy 80 GPISLLIIGIVAVHGMILVKCAHFCRLNKSFDYGDVTVMYGLSPCSWLNRNHAHWG 139
Db 82 GPLSLLVWGLIACHMHLVRCAQRFCHRLNKPMDYGDVTVMHGLAFSPNWLQNHAWG 141
Qy 140 RRVVDFFLIVTQGFCCVYFVFLADNPKQVTEAANGTNNCHNETVLTPTMDSRLYML 199
Db 142 RRVVSFFLIVTQGFCCVYFVFLADNLKQVVEAVNSTISCHKNETVLTPTMDSRLYML 201
Qy 200 SFLPFLVLLVFNRLRALSIFSLANITMVLVSLVMYQFIVQRIIDPSPHLPVAPWKTY 259
Db 202 SFLPVLGLLVFNRLRALTIFSLANISMLVSLVIAQYIIQEIIPDASQLPLVASWKTY 261
Qy 260 LFFGTAFSPGIGMVLPLENKMCKPRKFPILLYLGMVITVILYISLGLYLOFGANIQ 319
Db 262 LFFGTAFSPGIGMVLPLENKMCKDARFPITLSLGMSTITLYIAIGALGYLRFDDIK 321
Qy 320 GSITLNLPCNWLQSVKLLYIGIFFTYALQFYVPAEIIIPFVSRAPEHCELVDLFR 379
Db 322 ASITLNLPCNWLQSVKLLYVVGILCTYALQFYVPAEIIIPLASQVSKRWALPVDLSIR 381
Qy 380 TVLVCLTCILAILPRDLVLSVGSVSSSALALIIPLEVTTFYSGMSPLTIFKDAL 439
Db 382 LALVCLTCMLAILPRDLVLSVGSVSSSALALIIPLELVTVYVYGGISPLATVKDAL 441
Qy 440 ISILGFGVGVVGTVEALYELIQTSPNAPFINST 472
Db 442 ISILGFGVGVVGTVEALDELKSGNSPALSNST 474

RESULT 12
Q6NRA6_XENLA
ID Q6NRA6_XENLA PRELIMINARY; PRT; 479 AA.
AC Q6NRA6;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE MGC84608 protein.
GN Name=MGC84608;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Oocytes;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
```


Db 325 ASITLNLPCWLYQSVKLLYVVGILCTHALQFYVPAETIIIPLAVSQVSKRWALPVDLSIR 384
Qy 380 TVLVCLTCLAILPRDLVLISVGSVSSSALALIIPPLLEVTTFYSGMSPLTFKDAL 439
Db 385 LALVCTVCLAILPRDLVLISVGSVSSSALALIIPPLLEVTTFYSGMSPLTFKDAL 444
Qy 440 ISILGFGVGVGTVEALVELYELIQSNAPFINST 472
Db 445 ISILGFGVGVGTVEALVELYELIQSNAPFINST 477
RESULT 14
Q86YK4 HUMAN
ID Q86YK4 HUMAN PRELIMINARY; PRT; 313 AA.
AC Q86YK4;
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 11.
DE Transmembrane transport protein (Fragment).
GN Name=SLC36A1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15058382; DOI=10.1007/s00335-003-2319-3;
RA Birmingham J.R. Jr., Pennington J.;
RT "Organization and expression of the SLC36 cluster of amino acid
RT transporter genes.";
RL Mamm. Genome 15:114-125 (2004).
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC EMBL; AY227112; AAC37092.1; -; mRNA.
DR Ensembl; ENSG00000123643; Homo sapiens.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005279; P:amino acid-polyamine transporter activity; IEA.
DR GO; GO:0006865; P:amino acid transport; IEA.
DR InterPro; IPR002422; AA/rel_permease2.
DR InterPro; IPR013057; AA/transp_tm.
DR Pfam; PF01490; Aa_trans; 1.
KW Transmembrane.
FT NON TER
SQ -SEQUENCE 313 AA; 34934 MW; EFC510827A574E1C CRC64;
Query Match
Best Local Similarity 64.5%; Score 1592; DB 2; Length 313;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 164 DNFKQVIEAANGTTNNCHNNEVTILPTPMDSRLYMSLPELVLLVFTIRNLRAISIFSL 223
Db 1 DNFKQVIEAANGTTNNCHNNEVTILPTPMDSRLYMSLPELVLLVFTIRNLRAISIFSL 60
Qy 224 ANITMLSLVMYQFIVQRIIDPPSHLPLVAPWKYPLPFGTAIFSFEIGVMVPLENKKM 283
Db 61 ANITMLSLVMYQFIVQRIIDPPSHLPLVAPWKYPLPFGTAIFSFEIGVMVPLENKKM 120
Qy 284 DPRKPLILYUGMIVTILYISLGLCYLQFGANIQGSITLNLPCWLYQSVKLLYSIGI 343
Db 121 DPRKPLILYUGMIVTILYISLGLCYLQFGANIQGSITLNLPCWLYQSVKLLYSIGI 180
Qy 344 FFFVALQFYVPAEIIIPFVSVRAPEHCELVVDLFFVTVLVCLTCLAILPRDLVLISLV 403
Db 181 FFFVALQFYVPAEIIIPFVSVRAPEHCELVVDLFFVTVLVCLTCLAILPRDLVLISLV 240
Qy 404 GSVSSSALALIIPPLLEVTTFYSGMSPLTFKDALISILGFGVGVGTVEALVELYELIQPS 463
Db 241 GSVSSSALALIIPPLLEVTTFYSGMSPLTFKDALISILGFGVGVGTVEALVELYELIQPS 300

Qy 464 NAPIFINSTCAFI 476
Db 301 NAPIFINSTCAFI 313
RESULT 15
Q726B4 HUMAN
ID Q726B4 HUMAN PRELIMINARY; PRT; 470 AA.
AC Q726B4;
DT 01-OCT-2003, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2003, sequence version 1.
DT 07-FEB-2006, entry version 17.
DE Proton/amino acid transporter 3 (Solute carrier family 36
DE (Proton/amino acid symporter), member 3).
GN Name=SLC36A3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=22694810; PubMed=12809675; DOI=10.1016/S0888-7543(03)00099-5;
RA Boll M., Foltz M., Rubio-Alfaga I., Daniel H.;
RT "A cluster of proton/amino acid transporter genes in the human and
RT mouse genomes.";
RL Genomics 82:47-56 (2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX TISSUE=PCR rescued clones;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX TISSUE=PCR rescued clones;
RG NIH MGC Project;
RL Submitted (AUG-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC EMBL; AY162215; AAC11789.1; -; mRNA.
DR EMBL; BC101095; AAC101096.1; -; mRNA.
DR Ensembl; ENSG00000186334; Homo sapiens.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005279; P:amino acid-polyamine transporter activity; IEA.
DR GO; GO:0006865; P:amino acid transport; IEA.
DR InterPro; IPR002422; AA/rel_permease2.
DR InterPro; IPR013057; AA/transp_tm.
DR Pfam; PF01490; Aa_trans; 1.
SQ SEQUENCE 470 AA; 51735 MW; AA192584E6BFE848 CRC64;
Query Match 59.8%; Score 1476.5; DB 2; Length 470;

Best Local Similarity 62.2%; Pred. No. 9.5e-99;				
Matches 291; Conservative 77; Mismatches 95; Indels 5; Gaps 5;				
Qy	6	LRNEDYH-DYSSSTDVSP	EESSPSEGLNNLSPPGSYQ	REGQSNSTTWFTLIHLKGNIGTG 64
Db	3	LLGRDYNSELSLNDNGP-QSP	SSSSSITSSENVHPA-GEA-GLSM	QOTLIHLKGNIGTG 59
Qy	65	LLGLPLAVKNAIGVMGPI	SLIIIGIVAVCHMGILVCKAHHFCRR	LKNSFVDYGDVTMYGL 124
Db	60	LLGLPLAIKNAAGLVGV	SLLAIGLVTVHCVMLLNCAQHLSQR	LQKTFVNYGEATMYGL 119
Qy	125	ESSPCSWLRNHAHWGRV	DPFLIVTOLGFC	CVYFVLADNFKQVIEAANGTTNNCHNE 184
Db	120	ETCPNTWLRHAHVNGRY	TVTSFLLVITOLGFC	SVYFVFEMADNLQOMVEKAHVTSNICOPRE 179
Qy	185	TVILTPMTDSRLYMLSP	LPFLVLLVFI	RNRALSIFSLANITMLVSLVMIYOFIVQRIIP 244
Db	180	ILTLTPILDIRFYMLI	ILPFLILLVFIQNLKVL	SVFSTLANITTLGSMALIFEYIMEGIP 239
Qy	245	DPSHLPLVAPWKTYPL	FFGTAFISFEGIGMVLPLENKM	KDPRKPLILYLGMVIVTILYI 304
Db	240	YPSNLPLMANWKIFLL	FFGTAFITFEGVGMVPLKNQ	KHPQQFSFVLYLGMSIVILYI 299
Qy	305	SLGCLGYLQFGANIQ	GSITLNLPCWLYQSVKLLYSIG	IFFTYALQFYVPAEIIIPFVS 364
Db	300	LLGLTGYMKFGSDTQ	ASITLNLPCWLYQSVKLMYSIG	IFFTYALQFHPAEIIIPFAIS 359
Qy	365	RAPHECELVDL	LFVRTVLVCI	TCILATLIPRLDIVISLVGSVSSSALALIIIPPLEVTF 424
Db	360	QVSESWALFV	DLVSRSALVCLTC	VSALLIPRLDIVISLVGSVSSSALALIIIPALLEIVIF 419
Qy	425	YSEGMSPLTIFK	DALISILGFGVGVGT	YEALYELIQPSNAPIFINST 472
Db	420	YSEDMSCVTIAK	IMISIVGLLCIFG	TYQALYELPOPISHSM-ANST 466

Search completed: August 24, 2006, 01:24:53
Job time : 306 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2006, 01:25:09 ; Search time 50 Seconds
(without alignments)
833.292 Million cell updates/sec

Title: US-10-679-362-2
Perfect score: 2469
Sequence: 1 MSTQRLNEDYHDYSSTDVS.....YELIQSNAPFINSTCAFI 476

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA.*

- 1: /EMC_Celerra_SIDS3/ptodata/2/iaa/5_COMB.pep.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/iaa/6_COMB.pep.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/iaa/7_COMB.pep.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/iaa/H_COMB.pep.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/iaa/PTUS_COMB.pep.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/iaa/RE_COMB.pep.*
- 7: /EMC_Celerra_SIDS3/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2217	89.8	434	2	US-10-094-749-3076
2	1293.5	52.4	500	2	US-09-178-093B-26
3	1255.5	50.9	485	2	US-10-154-419-92
4	251.5	10.2	486	2	US-09-178-093B-1
5	229.5	9.3	525	2	US-09-178-093B-2
6	215.5	8.7	485	1	US-08-362-512A-2
7	215.5	8.7	485	2	US-08-964-939-2
8	215.5	8.7	485	2	US-09-854-774-2
9	203.5	8.2	108	2	US-09-599-360B-109
10	203	8.2	399	2	US-09-248-796A-20813
11	202.5	8.2	443	2	US-09-248-796A-20815
12	197	8.0	447	2	US-09-370-253-10
13	196	7.9	150	2	US-09-270-767-43308
14	187.5	7.6	447	2	US-09-370-253-6
15	187	7.6	449	2	US-09-640-419C-24
16	185.5	7.5	148	2	US-09-270-767-32811
17	185.5	7.5	148	2	US-09-270-767-48028
18	180.5	7.3	418	2	US-09-640-419C-25
19	179.5	7.3	493	1	US-08-362-512A-4
20	179.5	7.3	493	2	US-08-964-939-4
21	179.5	7.3	493	2	US-09-854-774-4
22	178.5	7.2	456	2	US-09-976-594-594
23	173.5	7.0	446	2	US-09-640-419C-26
24	173	7.0	287	2	US-09-248-796A-20822
25	171.5	6.9	452	2	US-09-640-419C-5
26	170	6.9	412	2	US-09-311-021-42

27	169	6.8	136	2	US-09-270-767-31761	Sequence 31761, A
28	169	6.8	136	2	US-09-270-767-46978	Sequence 46978, A
29	169	6.8	447	2	US-09-370-253-2	Sequence 2, Appl
30	166.5	6.7	462	2	US-09-640-419C-27	Sequence 27, Appl
31	165	6.7	504	2	US-09-919-497-67	Sequence 67, Appl
32	165	6.7	505	2	US-09-949-016-11556	Sequence 11556, A
33	164	6.6	504	2	US-09-178-093B-28	Sequence 28, Appl
34	162.5	6.6	134	2	US-09-270-767-42974	Sequence 42974, A
35	160.5	6.5	511	2	US-09-248-796A-20816	Sequence 20816, A
36	157.5	6.4	90	2	US-09-178-093B-37	Sequence 37, Appl
37	140.5	5.7	154	2	US-09-270-767-31638	Sequence 31638, A
38	140.5	5.7	154	2	US-09-270-767-46855	Sequence 46855, A
39	139.5	5.7	241	2	US-09-270-767-43616	Sequence 43616, A
40	138.5	5.6	90	2	US-09-178-093B-40	Sequence 40, Appl
41	136	5.5	415	2	US-09-602-787A-530	Sequence 530, Appl
42	133	5.4	116	2	US-09-270-767-32714	Sequence 32714, A
43	133	5.4	116	2	US-09-270-767-47931	Sequence 47931, A
44	127	5.1	259	2	US-09-370-253-4	Sequence 4, Appl
45	122.5	5.0	463	2	US-09-583-110-5223	Sequence 5223, Ap

ALIGNMENTS

RESULT 1
US-10-094-749-3076
; Sequence 3076, Application US/10094749
; Patent No.-6979557
; GENERAL INFORMATION:

; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YANAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, KEIOPARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOKYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3076

; LENGTH: 434
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-3076

Query Match 89.8%; Score 2217; DB 2; Length 434;
Best Local Similarity 90.8%; Pred. No. 2.4e-221;
Matches 432; Conservative 1; Mismatches 1; Indels 42; Gaps 1;
QY 1 MSTQRLNEDYHDYSSTDVSPEESSEGNNLSSPGSYQRFQGSNSTTWFTLIHLKGN 60
DB 1 MSTQRLNEDYHDYSSTDVSPEESSEGNNLSSPGSYQRFQGSNSTTWFTLIHLKGN 60
QY 61 IGTGLLGLPLAVKNAIGIVMGPIISLIIGIVAVHCVKILVKAHFCRLNKSFDYGDTV 120

```
Db 61 IGTLGLPLAAKNAGIWMGPISLLIIGIVAVHGMGILVKAHHFCRRLNKSFVDVGDV 120
Qy 121 MYGLESPCSWLRNHAHWGRVVDFFLIIVTQLGFCVCYFVFLADNFKQVIEAANGTTNCC 180
Db 121 MYGLESPCSWLRNHAHWGRVVDFFLIIVTQLGFCVCYFVFLADNFKQVIEAANGTTNCC 180
Qy 181 HNNETVILTTMDSRLYMLSFPLVLLVFIIRNLALSIFSLLANITMLVSLVMYQFIV 240
Db 181 HNNETVILTTMDSRLYMLSFPLVLLVFIIRNLALSIFSLLANITMLVSLVMYQFIV 240
Qy 241 QRIIPDPHPLVAPWKTYPLPFGTAIFSPGIGMWLPLENKMDPRKFPPLILYGMVIT 300
Db 241 QRIIPDPHPLVAPWKTYPLPFGTAIFSPGIGMWLPLENKMDPRKFPPLILYGMVIT 300
Qy 301 ILXISLGLGYLOFGANIQGSITLNLPCNWLQSVKLLYSIGIFFYFVALQFYVPAEIIIP 360
Db 301 ILXISLGLGYLOFGANIQGSITLNLPCNWLQSVKLLYSIGIFFYFVALQFYVPAEIIIP 360
Qy 361 FFVSRAPHELCVVDLFRVTVLVCLTCLAILPRDLVLSLGVSSSSALALIIPPLE 420
Db 361 FFVSRAPHELCVVDLFRVTVLVCLTCLAILPRDLVLSLGVSSSSALALIIPPLE 420
Qy 421 VTFYSEGMSPITFKDALISILGFVGVVGTVEALYELIQSNAPIFINSTCAFI 476
Db 421 VTFYSEGMSPITFKDALISILGFVGVVGTVEALYELIQSNAPIFINSTCAFI 476
Qy 476 VTFYSEGMSPITFKDALISILGFVGVVGTVEALYELIQSNAPIFINSTCAFI 476
Db 476 VTFYSEGMSPITFKDALISILGFVGVVGTVEALYELIQSNAPIFINSTCAFI 476
RESULT 2
US-09-178-093B-26
; Sequence 26, Application US/09178093B
; Patent No. 6660846
; GENERAL INFORMATION:
; APPLICANT: Robert H. Edwards
; APPLICANT: Richard J. Reimer
; APPLICANT: Steve L. McIntire
; APPLICANT: Erik M. Jorgenson
; APPLICANT: Kim Schuske
; TITLE OF INVENTION: Vesicular Amino Acid Transported
; FILE REFERENCE: 2002-0005.30
; CURRENT APPLICATION NUMBER: US/09/178,093B
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/063,012
; PRIOR FILING DATE: 1997-10-23
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide (RVT2 polypeptide sequence)
US-09-178-093B-26
```

```
Query Match 52.4%; Score 1293.5; DB 2; Length 500;
Best Local Similarity 54.7%; Pred. No. 2.2e-125;
Matches 263; Conservative 80; Mismatches 117; Indels 21; Gaps 6;

Qy 6 LRNEDYHDSSTDVSPSESPSEGLNNLSSPGSYQRFQGSNSTTWFTQLIHLKKNIGTGL 65
Db 23 LINEQNFDGSS-----DEEQEQLT--LPMQKHVQLDQ--HGISFVQTLMLHLLKKNIGTGL 74
Qy 66 LGPLAVKNAAGIWMGPISLLIIGIVAVHGMGILVKAHHFCRRLNKSFVDVGTVMYGL 125
Db 75 LGPLAIKNAAGIVLGPISLVFIIGIIVHGMHILVRCSHFLCQPKFKSTGLCYSTVSFAME 134
Qy 126 SSPCSWLRNHAHWGRVVDFFLIIVTQLGFCVCYFVFLADNFKQVIEAANGTT---NCHN 182
Db 135 ASPWSCLOQAAWGRSVVDFFLIIVTQLGFCVCYFVFLADNFKQVIEAANGTT---NCHN 194
Qy 183 NETVILTTMDSRLYMLSFPLVLLVFIIRNLALSIFSLLANITMLVSLVMYQFIV 242
Db 195 LSQVCERRSVDLRYMLCFLPILLVFIIRNLALSIFSLLANITMLVSLVMYQFIV 254
```

```
Qy 243 IPDPHPLVAPWKTYPLPFGTAIFSPGIGMWLPLENKMDPRKFPPLILYGMVITVL 302
Db 255 MPDPHNLPIVAGWKYPLPFGTAIFSPGIGMWLPLENKMDPRKFPPLILYGMVITVL 314
Qy 303 YISLGLGYLOFGANIQGSITLNLPCNWLQSVKLLYSIGIFFYFVALQFYVPAEIIIPF 361
Db 315 YISLATLGYMCFRDEIKGSITLNLPCNWLQSVKLLYSIGIFFYFVALQFYVPAEIIIPA 374
Qy 362 FVSRAPEHCELVVDLFRVTVLVCLTCLAILPRDLVLSLGVSSSSALALIIPPLEV 421
Db 375 VVARLHAKWKICDFGIRSLVITCAGAVLIIRLDIVISFVCAVSSSTLALILPPLVEI 434
Qy 422 TTFYSEGMSPITFKDALISILGFVGVVGTVEALYELIQSNAPIFINST 472
Db 435 LTFSKDHYNNVMVKNISIAFTGFVGLLGTYYTVEEIIYPTTAVADGASQSLSLVNST 494
Qy 473 C 473
Db 495 C 495

RESULT 3
US-10-154-419-92
; Sequence 92, Application US/10154419
; Patent No. 6972187
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Gluckman, Maria Alexandra
; APPLICANT: Meyers, Rachel E.
; TITLE OF INVENTION: NOVEL 38594, 57312, 53659, 57250, 63760, 49938, 32146,
; TITLE OF INVENTION: 57259, 67118, 67067, 62092, 8099, 46455, 54414, 53763,
; TITLE OF INVENTION: 67076, 67102, 44181, 67084FL, 67084ALT, FBH58235FL, 57255,
; TITLE OF INVENTION: AND 57255alt MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-249
; CURRENT APPLICATION NUMBER: US/10/154,419
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: Fast-SEQ Version 4.0
; SEQ ID NO 92
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-154-419-92
```

```
Query Match 50.9%; Score 1255.5; DB 2; Length 485;
Best Local Similarity 52.0%; Pred. No. 1.8e-121;
Matches 254; Conservative 80; Mismatches 119; Indels 35; Gaps 6;

Qy 6 LRNEDYHDSSTDVSPSESPSEGLNNLSSPGSYQRFQGSNSTTWFTQLIHLKKNIGTGL 65
Db 7 LINEQNFDGTSDEEHEQE-----LLPVQKHVQLDDQEGISFVQTLMLHLLKKNIGTGL 58
Qy 66 LGPLAVKNAAGIWMGPISLLIIGIVAVHGMGILVKAHHFCRRLNKSFVDVGTVMYGL 125
Db 59 LGPLAIKNAAGIVLGPISLVFIIGIIVHGMHILVRCSHFLCFLFKKSTGLCYSTVSFAME 118
Qy 126 SSPCSWLRNHAHWGRVVDFFLIIVTQLGFCVCYFVFLADNFKOV-----IEAANG 175
Db 119 VSPWSCLOQAAWGRSVVDFFLIIVTQLGFCVCYFVFLADNFKOV-----IEAANG 178
Qy 176 TTNCHNNETVILTTMDSRLYMLSFPLVLLVFIIRNLALSIFSLLANITMLVSLVMI 235
Db 179 SSNPCEER-----SVDLRIYMLCFLPILLVFIIRNLALSIFSLLANITMLVSLVMI 231
Qy 236 YQFIVQRIIDPSPHPLVAPWKTYPLPFGTAIFSPGIGMWLPLENKMDPRKFPPLIYL 295
Db 232 YQVVRNPDHNLPIVAGWKYPLPFGTAIFSPGIGMWLPLENKMDPRKFPPLIYL 291
Qy 296 MIVTILYSLGLCYLOFGANIQGSITLNLPCNWLQSVKLLYSIGIFFYFVALQFYV 354
Db 292 MGIIVTILYSLGLCYLOFGANIQGSITLNLPCNWLQSVKLLYSIGIFFYFVALQFYV 351
```

Qy 355 AEIIIPFVSRAPEHCELVVDLFRVTVLVCITLAILIPLDLVLSVGVSSSALALI 414
 Db 352 AEIIIPGITSFHTKQWQICFGIRSLVSTCAGAILIPDLVIFSVGAVSSSTLALI 411
 Qy 415 IPPLEVTTFYSEGMSPITIFKDALISILGFGVGVGYEALYELIQPS-----NAP 466
 Db 412 LPPLVEILTFSEKHYNIMWLKNISIAFTGVVGLLGTITVBEIIVPTPKVAVGTQSP 471
 Qy 467 IF-INSTC 473
 Db 472 FLNLNSTC 479

RESULT 4
 US-09-178-093B-1
 ; Sequence 1, Application US/09178093B
 ; Patent No. 6660846
 ; GENERAL INFORMATION:
 ; APPLICANT: Robert H. Edwards
 ; APPLICANT: Richard J. Reimer
 ; APPLICANT: Steve L. McIntire
 ; APPLICANT: Erik M. Jorgenson
 ; APPLICANT: Kim Schuske
 ; TITLE OF INVENTION: Vesicular Amino Acid Transported
 ; FILE OF INVENTION: Composition and Method
 ; FILE REFERENCE: 2002-0005.30
 ; CURRENT APPLICATION NUMBER: US/09/178,093B
 ; CURRENT FILING DATE: 2001-08-20
 ; PRIOR APPLICATION NUMBER: 60/063,012
 ; PRIOR FILING DATE: 1997-10-23
 ; NUMBER OF SEQ ID NOS: 50
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 486
 ; TYPE: PRT
 ; ORGANISM: Caenorhabditis elegans
 ; FEATURE:
 ; NAME/KEY: CHAIN
 ; LOCATION: (1)...(486)
 ; OTHER INFORMATION: UNC-47 polypeptide sequence
 ; US-09-178-093B-1

Query Match 10.2%; Score 251.5; DB 2; Length 486;
 Best Local Similarity 21.7%; Pred. No. 3.1e-17;
 Matches 107; Conservative 83; Mismatches 171; Indels 131; Gaps 18;
 Qy 22 EESPSEGLNNLSSPGSYQRFQCSNSTTWFTLIHLKNGNIGTGLGLPLAVK-----NAG 76
 Db 75 BEAKDDGHGEASEPISA-----LQAANVTNATQGMFIVGLPIAVKVGWWSIG 123
 Qy 77 IVMGPISLIIGIVAVHCGM---ILVKCAHFRCRLNKSFDYDGTVMYGLSSPCSWLR 133
 Db 124 AMVG-----VAYVCYWTGVLLIECLYENGKVKRKY----- 154
 Qy 134 NHAHGRVRVDFF-----LVITV---LGFCCVYFVFLADNPKQVTEAANGTNNCHN 182
 Db 155 -----REIADYKPGFKWVLAQAQLTETLCIYLVLAAD-----LLQSCF- 196
 Qy 183 NETVILPTMDSRLYMLSPFLVLLVFIIRNLRLSIFSLANIT-MLVSLVMIYQFTVQ 241
 Db 197 -----PSVDKAGMMITSAULTCSFLDDLQIVSRLSFFNAISHLVNLIMVLYC--- 246
 Qy 242 RIPDPSHLPLVAPW-----KTYPLFFGTAFISFEGIGMVLPLENKKDKRPFPL 290
 Db 247 -----LSFVSQMSFSTITFSLNITLPTIVGMVVFYGTSHIFLPNLEGNMKNPAQFN 299
 Qy 291 ILYLGMVITVILYISLGCIGYLOFGANIQQSITLNLNPNCLWYQSVKLLYS-----IGIFFT 346
 Db 300 MLKWSHIAAAVFKVVGMLGFTFGELTQBEISNLPN-----QSPKILVNLILVVKALLS 355
 Qy 347 YALQFYVPAEII-----IPFVSRAPHEHCELVVDLFRVTVLVCITLAILIPLRL 396

Db 356 YPLFPYAAVQLKNNLFLGYPQTPFTSCYSPDKSLREWAVTLRIILVLTFLVALSVPYL 415
 Qy 397 DLVLSVGVSSSALALIIPPLEVTTTFYSEGMSPITIFK---DALISILGFGVGVGY 453
 Db 416 VELMGLVGNITGTMLSFIWPAFLH--YIKE--KTLANNFEKRFDOGIIMGCSVCISGVY 471
 Qy 454 EALYELIQPSNA 465
 Db 472 FSSMELLRAINS 483

RESULT 5
 US-09-178-093B-2
 ; Sequence 2, Application US/09178093B
 ; Patent No. 6660846
 ; GENERAL INFORMATION:
 ; APPLICANT: Robert H. Edwards
 ; APPLICANT: Richard J. Reimer
 ; APPLICANT: Steve L. McIntire
 ; APPLICANT: Erik M. Jorgenson
 ; APPLICANT: Kim Schuske
 ; TITLE OF INVENTION: Vesicular Amino Acid Transported
 ; FILE OF INVENTION: Composition and Method
 ; FILE REFERENCE: 2002-0005.30
 ; CURRENT APPLICATION NUMBER: US/09/178,093B
 ; CURRENT FILING DATE: 2001-08-20
 ; PRIOR APPLICATION NUMBER: 60/063,012
 ; PRIOR FILING DATE: 1997-10-23
 ; NUMBER OF SEQ ID NOS: 50
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 525
 ; TYPE: PRT
 ; ORGANISM: Rattus norvegicus
 ; FEATURE:
 ; NAME/KEY: CHAIN
 ; LOCATION: (1)...(525)
 ; OTHER INFORMATION: RUNC-47 polypeptide sequence
 ; US-09-178-093B-2

Query Match 9.3%; Score 229.5; DB 2; Length 525;
 Best Local Similarity 22.0%; Pred. No. 6.6e-15;
 Matches 95; Conservative 82; Mismatches 178; Indels 77; Gaps 15;
 Qy 62 GTGLGLPLAVKNAIGVMGPISLLIIGIVAVHCGMLVKAHFRCR-----RLNKSFFDY 116
 Db 131 GMFVLGFLPYAILHGG-VLGLFLIIFAAVCCYTGKILLIACLYEENEDEGEVVRDYSV-- 187
 Qy 117 GDTVMYGLSSPCSWLRNHAHWRVRVDFFLIIVTQLGCCVYFVFLADNPKQVTEAANGT 176
 Db 188 -----AIANACCA--PRFPTLGGRVNVVAQII-ELVMTCILYVVVSGNL----- 228
 Qy 177 TNNCHNETVILPTMDSRLYMLSPFLVLL--VFIRNLRLSIFSLANITMLVSLVM 234
 Db 229 ---MYNS-----PFLPVQSQKSWSIATAVLLPCAFLEKNKAVSKFSLCLTAHFVNIL 280
 Qy 235 IYQIVORIPDPSHLPLVAPW-----KTYPLFFGTAFISFEGIGMVLPLENKK 283
 Db 281 VIAYCLGRAD-----WAEKVKPYIDVKFPISIGIIVFSYTSQIFLPSLEGNMQ 331
 Qy 284 DPRKFPILYLGMMVITVILYISLGCIGYLOFGANIQQSITLNLNPNCLWYQSVKLLYSIGI 343
 Db 332 QPSEFHCMMNWTIAACVCLKGLFALVAYLTWADETKEVITDNLPG-----SIRAVVNIFL 386
 Qy 344 ----FFTYALQFYVPAEII-----IPFVSRAPHEHCELVV-DLFRVTVLVCITCIL 389
 Db 387 VAKALLSYPLPFFAAVEVLEKSLFQEGSRAPFPACYGGDGRKMGWGLTLCALVVFVTLIM 446
 Qy 390 AILPLDLVLSVGVSSSALALIIPPLEVTTTFYSEGMSPITIFKDALISILGFGVGV 449
 Db 447 AIYPHPFALMGLTGTGAGLCFLPSLFLRLLRK-LLWHQVFFDVAIFVIGICSV 505
 Qy 450 VGTVEALYELIQ 461

Db 60 LAWAIAQLGWIAGTSILLIFSFITVFTSTMLADCYRAPDPVTGKRNVTYMDVRSYLGGR 119
Qy 125 ESSPCSWLRNHAHGRVDFLLVTLVQGLFCVVFVFLADNFKQVIEAANGTTNNCHNE 184
Db 120 KVQLC-----GVAQYGN-----LIGVTGYTITASISLVAVGK-----SNCFHDK 159
Qy 185 TVILPTMDSRLYMLSPFLPVLVLLVFNRLRALSIFSLANI-----FGTAIFSPG 271
Db 160 GHTADCTISNYPYMAVFGIIQVILSQIPNPKLSFLSIMAAVMSFTYATIGIGLAIAATVA 219
Qy 227 -----TMLVSLVMVYQFIVQRIQVDPDPSHLPVAPWKTYPLF-----FGTAIFSPG 271
Db 220 GGVKGTSMGTAVGVDDVTAQAQKI-----WRSFQAVGDIAPAYATVLIETQD 268
Qy 272 IGMVLPLENKMDRKPFLILYLGWIVITILYISLGLCYLQFGANIGSITLN---LPN 328
Db 269 TLRSSPAENKAM--KRASLV---GVSTTFYILCGCIGYAAFGNNAAPGDFLTDGFPFEP 323
Qy 329 CWLYQ-----SVKLYSIGIFFTYALQF-----YVPAEIIIPFVSRAPEHCELVV 374
Db 324 FWLIDFANACIAVHLIGAYQVFAQPIQFQVEKKCRNRYPDNKFITSEYSVNVVPLGKFN 383
Qy 375 DLF---VRLVLCVTCILAILPRLDLVISLVGSSSALALIIPLLEVTTFYSEGMS 431
Db 384 SLFRLVWRTAYVVTVVAMIFPFNAILGLIGAASFPLTVYPPVEMHIAQTKIKKYS 443
Qy 432 LTI-----FKDALISILGFGVGVV 451
Db 444 RWIAKTMCYVCLIVSLAAGSIAG 469

RESULT 8
US-09-854-774-2
; Sequence 2, Application US/09854774
; Patent No. 6809233
; GENERAL INFORMATION:
; APPLICANT: FROMMER, Wolf-Beird
; TITLE OF INVENTION: DNA SEQUENCES FOR AN AMINO ACID
; TRANSPORTER, PLASMIDS, BACTERIA, YEASTS AND PLANTS
; CONTAINING A TRANSPORTER AND THEIR USE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen
; STREET: 1180 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10036-8403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/854,774
; FILING DATE: 14-May-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/964,939
; FILING DATE: 1998-05-28
; APPLICATION NUMBER: US 08/007,636
; FILING DATE: 21-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Meilman, Edward A.
; REGISTRATION NUMBER: 24,735
; REFERENCE/DOCKET NUMBER: P/951-107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 382-0700
; TELEFAX: (212) 382-0888
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 485 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-854-774-2

Query Match 8.7%; Score 215.5; DB 2; Length 485;
Best Local Similarity 20.0%; Pred. No. 1.7e-13;
Matches 101; Conservative 81; Mismatches 221; Indels 103; Gaps 15;

Qy 8 NEDYHVSSTDSPEESPSEGLNNLSSPGSVQRQGSNSTTFQTLHLLKGNIGTGLG 67
Db 5 NTEGHNSTAESDAYTVSDPTKNVDEDEKRTG-----TWLTASAHIIITAVIGSGVLS 59
Qy 68 LPLAVKNAGIVMGPISLIIGIVAVHGMGILVKC---AHHFCRRLNKSFVDYGDVTMYGL 124
Db 60 LAWAIAQLGWIAGTSILLIFSFITVFTSTMLADCYRAPDPVTGKRNVTYMDVRSYLGGR 119
Qy 125 ESSPCSWLRNHAHGRVDFLLVTLVQGLFCVVFVFLADNFKQVIEAANGTTNNCHNE 184
Db 120 KVQLC-----GVAQYGN-----LIGVTGYTITASISLVAVGK-----SNCFHDK 159
Qy 185 TVILPTMDSRLYMLSPFLPVLVLLVFNRLRALSIFSLANI-----FGTAIFSPG 271
Db 160 GHTADCTISNYPYMAVFGIIQVILSQIPNPKLSFLSIMAAVMSFTYATIGIGLAIAATVA 219
Qy 227 -----TMLVSLVMVYQFIVQRIQVDPDPSHLPVAPWKTYPLF-----FGTAIFSPG 271
Db 220 GGVKGTSMGTAVGVDDVTAQAQKI-----WRSFQAVGDIAPAYATVLIETQD 268
Qy 272 IGMVLPLENKMDRKPFLILYLGWIVITILYISLGLCYLQFGANIGSITLN---LPN 328
Db 269 TLRSSPAENKAM--KRASLV---GVSTTFYILCGCIGYAAFGNNAAPGDFLTDGFPFEP 323
Qy 329 CWLYQ-----SVKLYSIGIFFTYALQF-----YVPAEIIIPFVSRAPEHCELVV 374
Db 324 FWLIDFANACIAVHLIGAYQVFAQPIQFQVEKKCRNRYPDNKFITSEYSVNVVPLGKFN 383
Qy 375 DLF---VRLVLCVTCILAILPRLDLVISLVGSSSALALIIPLLEVTTFYSEGMS 431
Db 384 SLFRLVWRTAYVVTVVAMIFPFNAILGLIGAASFPLTVYPPVEMHIAQTKIKKYS 443
Qy 432 LTI-----FKDALISILGFGVGVV 451
Db 444 RWIAKTMCYVCLIVSLAAGSIAG 469

RESULT 9
US-09-599-3608-109
; Sequence 109, Application US/095993608
; Patent No. 6548633
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Bougueret, L.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
; FILE REFERENCE: GENSET.050CP3
; CURRENT APPLICATION NUMBER: US/09/599,3608
; CURRENT FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: 60/113,686
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/141,032
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 09/469,099
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: Patent.pm
; SEQ ID NO 109
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:

```
; NAME/KEY: SIGNAL
; LOCATION: -28.-1
US-09-599-360B-109

Query Match      8.2%  Score 203.5; DB 2; Length 108;
Best Local Similarity 48.4%; Pred. No. 3.4e-13;
Matches 45; Conservative 16; Mismatches 23; Indels 9; Gaps 2;

Qy 390 AILPRDLVLISVGVSSSALALIIPLLEVTTFYSEGMSPLTIKDALISILGVGVFV 449
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 10 AILPRDLVIVSFVGVSSSTLALILPLVEILTFSEKHYNIWMVLKNISIAFTGVVGFL 69

Qy 450 VGYEALYELIQPS-----NAPIF-INSTC 473
      :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 70 LGTVITVEEIIYPTKVVAGTQSPFLNLNSTC 102

RESULT 10
US-09-248-796A-20813
; Sequence 20813, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248.796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 20813
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-20813

Query Match      8.2%  Score 203; DB 2; Length 399;
Best Local Similarity 21.9%; Pred. No. 2.5e-12;
Matches 91; Conservative 66; Mismatches 145; Indels 114; Gaps 15;

Qy 54 IHLKGNIGTGLGLPLAVKNAGIVMGPISLIIIGIVAVHCMGILVKCAHFCRRLNKSF 113
      :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 18 ISLVKTTIAGLLSMLAYSTDGIIFGTFIILAAFTS--GFGFLQC-----Y 64

Qy 114 VDYGDTVMYGLESFPCSWLRNHAHWRVDFFLIVTQGFCCVYFVLADNFKQVIEAA 173
      :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 65 VSRVVPVKHATFFNLCSITYPHL---SVVDFDAIYAVQFCGCAVSYLVLIRDLMP 117

Qy 174 NGTTNNCHNETVILPTMDSRLYMLSF---PFLVLLVFIRNLRALSIFSLIANITML 229
      :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 118 -----VTYVPIDEKHYPLFWLLVSTVLTITPLSFLKMLDSLKYSIISLGLVAIF 165

Qy 230 -VSLVMVYQFIVQIPDFSHLPLVAPKTYPLF--FGTAIFSPFEGIGMVLPLENKKDPR 286
      :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 166 YMTILVIGHYFAGDIERQGOITLF-PTSVTGVPFTFSIIIVFAFTGHQWFSIINEARDKS 224

Qy 287 KFPPL--ILYLGWVITVILYISLGLCLGLOFGANTQGSITLNLPCNWLVSQVSLYSIGIF 344
      :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 225 LTSITKLNVNFAIGSSLLFLAVGLSGYLTFQGDVGNVILYPN-----GLTTTIGRF 277

Qy 345 -----FTVALQYVPAEI----- 357
      :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 278 CIVFMVTFSPFLMIH-PARISINNIYHWVTNYFDKDEANESTALLGNESQESQIAEDEE 336

Qy 358 -----IIPFVSRAPHEHCELVDLFLVTVLVCLTCILAILPRDLVLISLVGS 405
      :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 337 QGKTHLSHVVPF-----PHKHFI---VTTSLLIIGVLLAISIKSPALILAVGA 383

RESULT 11
```

```
US-09-248-796A-20815
; Sequence 20815, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248.796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 20815
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-20815

Query Match      8.2%  Score 202.5; DB 2; Length 483;
Best Local Similarity 21.3%; Pred. No. 3.7e-12;
Matches 108; Conservative 76; Mismatches 179; Indels 145; Gaps 20;

Qy 13 DYSSTDVSPESPSESGINNLSPPGSYQRFQGS---NSTTWFOQLIHLKGNIGTGLGLP 69
      :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 51 DWSNNSVFDE--TSELLPTLSRV-STKRHSFSLITGNSTAAQTIFNSINTLIGTGLMLSLP 107

Qy 70 LAVKNAGIVMGPISLIIIGIVAVHCMGILVKCAHFCRRLNKSFDYDGTVMYGLESPPC 129
      :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 108 LGFKMSGWLFSGLSLLVVSAPLNTTAKYLGKILYR-----HQELMTYGDIA----- 153

Qy 130 SWLRNHAHWR---RVVDFFLIIVTQGFCCVYFVLADNF-----KQVIEAANGT 176
      :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 154 -----YAYGGKPYLVLTLPFFVIDLFGASLTLLILFADSFITVWPHVPALKAIIVAV-- 205

Qy 177 TNNCHNETVILPTMDSRLYMLSFPLFVLLVFIRNLRALSIFSLIANITM--LVSLVM 234
      :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 206 -----VFVLSLLP-----LSMLSIFSLIIGIISTVGIILSVF 236

Qy 235 IYQFIVQIPDPShLP-----LVAPWKTYPLFFGTALFSEGIGMV 275
      :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 237 ICGLFVDTSPGSLIIPATTTLLPPNINLLFSLIGIFMAPGGHPVF----- 282

Qy 276 LPLENKKDPRKPEPLIYLGWVITVILYISLGLCLGLOFGANTQGSITLNL---PN--- 328
      :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 283 PELYRDMRHPFKETKSSNISFLVTLDDFSIGATGYLMYGLMVDDSIKVSIMQNPYPPI 342

Qy 329 -----CWLQYOSVKL---LYSIGIFTYALQFVVPABIIIPFFVSRAPHEHCELV-----V 374
      :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 343 INSILCILMGILPISKLPVTKPIITSYENIFGTAKYV-----KLDENGLTDTYGPT 396

Qy 375 DLFVRTVIVCLTCILAILPRDLVLISLVGSVSSSALALIIPLLEVTTFY-----SEG 428
      :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 397 RVFSTLSFCVLLISALLTTSFKGLVAFLSAICDVTVCLTLPLL-----FYLKLRSSVG 451

Qy 429 MSPLTIFFKDALISILGFVGVGVGTGYEAL 456
      :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 452 KLERLLIKIGI--VFSITATILGTYASI 477

RESULT 12
US-09-370-253-10
; Sequence 10, Application US/09370253
; Patent No. 6165792
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Sakai, Hajime
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: Amino Acid Transporters
; FILE REFERENCE: BB-1200
; CURRENT APPLICATION NUMBER: US/09/370,253
```

```
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: 60/097,222
; EARLIER FILING DATE: August 20, 1998
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-09-370-253-10

Query Match      8.0%; Score 197; DB 2; Length 447;
Best Local Similarity 21.3%; Pred. No. 1.2e-11;
Matches 93; Conservative 84; Mismatches 214; Indels 46; Gaps 15;

Qy 44 SNTTTFQTLHLKNGIGTGLGLPLAVKNAGIWMGPISLLIIGIVAHVCMGILVKCAH 103
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 33 SRNAKWYSAFHNTVMAGVLSLPYAMSELGPGIADVIVSWITLYTLQWVE-MH 91
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 104 HFCRLNKSFDYDGTVMYGLSESPCSWLRNHAHWRVDFFLIVTQLGFCVCYFVFLA 163
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 92 EMVP--GKRPDRYHELGHAFGDKLGLWI-----VVPQQLIV-EVGYNIVVMVTGG 139
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 164 DNFKQVTEAANGTTNNCHNETVILPTMDRLYMLSPFLPLVLLVFIIRNLALSIF 223
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 140 RSLKKFHDVI--CDGKCKDKILTFPIMIFASVHFVLSQLP-----NLNSISGVSLA 188
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 224 ANITMLVSLVMYQFIQVOR-----IPDPSHLPLVAPWKTYPLF--FGTAIFSPEGIMVLP 277
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 189 AAVMSLSYSTIANGASVDKGMVDVYNLRATTTGKRVGFGFALGEVAFAGHNVVLE 248
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 278 LE----NKKMDPRKFPLI--LYLGMVIVITILYISLGLGYLQFGANIQGS--ITLNLNPN 329
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 249 IQATIPSTPEKPSKKPMKGVVAVIYVVALCYFPVALLIGYAWFGNSVDDNLIITLNKPT- 307
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 330 WLYOSVKL---LYSIGIFFTVALQFYVPAEIIIPFFVSRAPHEHCELVVDLFRVTVLCLT 386
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 308 WLISTANMMVVIHVIGSYQIYAMPVDFMDIETVL---VKLKFPPGLTLRLIARTLYVAFT 364
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 387 CILAILPRLDLIVSLGVSSSALALIIPLEVTTFYSEGMSPLTIKDALISILGFV 446
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 365 MFVAITFPFGLGLGFGGFAPFTTYPFCIMLAIYKPRFS--LSWFTNWVCIVLGV 423
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 447 GFVVGTYEALYELIQPS 463
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 424 LMILSPIGGLRQIILDS 440
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
US-09-270-767-43308
; Sequence 43308, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 43308
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-43308

Query Match      7.9%; Score 196; DB 2; Length 150;
Best Local Similarity 35.5%; Pred. No. 3.2e-12;
Matches 43; Conservative 24; Mismatches 50; Indels 4; Gaps 3;

Qy 14 YSST-DVSPSPSEGLNLSPPGSYQRFQGSNTTFQTLIHLKNGNIGTGLGLPLAV 72
```

```
; Sequence 24, Application US/09640419C
; Patent No. 6630615
; GENERAL INFORMATION:
; APPLICANT: Bidney, Dennis L
; APPLICANT: Crasta, Oswald R
; APPLICANT: Hu, Xu
; APPLICANT: Lu, Guihua
; TITLE OF INVENTION: DEFENSE-RELATED SIGNALING GENES AND METHODS OF USE
; FILE REFERENCE: 35718/199009 (5718-92)
; CURRENT APPLICATION NUMBER: US/09/640,419C
; CURRENT FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/149,656
; PRIOR FILING DATE: 1999-08-18
; PRIOR APPLICATION NUMBER: 60/206,405
; PRIOR FILING DATE: 2000-05-23
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-640-419C-24

Query Match          7.6%; Score 187; DB 2; Length 449;
Best Local Similarity 20.7%; Pred. No. 1.4e-10;
Matches 82; Conservative 79; Mismatches 188; Indels 48; Gaps 15;

Qy 44 SNSTWFQTLJHLLKGNIGTGLGLPLAVKNAGIVMGPSISLLIIGIVAVHVMGILVKCAH 103
Db 28 SRNAKMYSAFHNTAMVAGVLSLPYAMNGLWGPCGVTIMVMSWIITLYTLQWQVE-MH 86

Qy 104 HFC--RRLNKSFDYDGTVMYGLESPSCSWLRNHAHMGRRVVDFFLIIVTQLGFCVCYVFV 161
Db 87 EIVPGKRLDR----YHELQGHAFGEKGLMI-----VVPQQLIV-EVGVDIVVMVT 132

Qy 162 LADNFKQVIEAANGTTNNCHNETVILPTWDSRLYMLSPFLPLVLLVFTIRNLRAISIFS 221
Db 133 GGASLKKVHQL---VCPDCKEIRTFWIMIFASVHFVISHLP-----NFNSISILS 180

Qy 222 LLANITMLVSLVMIYQIVOR--IPDPSHLPLVAP--WKTYPLF--PGTAIFSPEGIGMV 275
Db 181 LAAVMSLTYSTIAWAASVHKGVHPDVSFRASTDVGKVFNFNLALGDVAFAGHNVV 240

Qy 276 LPLENKMKD---PRKEPLI--LYLGMVIVITLYISLGLGYLQFGANIQGSITLNLPN- 328
Db 241 LEIQATIPSTPEMPSKVPMMRGVIVAVVVAICVFPVAFGLGYVFGNSVDDNIIITLEKP 300

Qy 329 CWLYQSCKLLYSIGIFTYALQFV-----PA-EIIPFPVSRAPCEHCELVVDLFRVTVL 382
Db 301 IWLIAAMNMFVIVHIGSYQVFHILIFAMFVDFMLETVLVKKMNFNPNPKLRFITRSLY 360

Qy 383 VCLTCILAILPLRLDLVLSLVSVSSSALALIIPPLL 419
Db 361 VAFTMIVAIQVPFPGGLGFGFGFAPAPTYYLPCLIM 397
```

Search completed: August 24, 2006, 01:26:34
Job time : 51 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2006, 01:25:54 ; Search time 185 Seconds
(without alignments)
1191.839 Million cell updates/sec

Title: US-10-679-362-2

Perfect score: 2469

Sequence: 1 MSTQRLRNEDYHYSSTDVS.....YELIQSNAPFINSTCAFI 476

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA Main:

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_PUBCOMB.pep:**
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_PUBCOMB.pep:**
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_PUBCOMB.pep:**
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10A_PUBCOMB.pep:**
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10B_PUBCOMB.pep:**
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_PUBCOMB.pep:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2469	100.0	476	US-09-805-456-2	Sequence 2, Appli
2	2469	100.0	476	US-10-225-810-21	Sequence 21, Appli
3	2469	100.0	476	US-10-679-362-2	Sequence 2, Appli
4	2460	99.6	496	US-10-080-334-22	Sequence 22, Appli
5	2402.5	97.3	568	US-10-332-447-8	Sequence 8, Appli
6	2398	97.1	500	US-10-080-334-20	Sequence 20, Appli
7	2324	94.1	476	US-09-805-456-7	Sequence 7, Appli
8	2324	94.1	476	US-10-679-362-7	Sequence 7, Appli
9	2217	89.8	434	US-10-094-749-3076	Sequence 3076, Ap
10	2217	89.8	434	US-10-080-334-137	Sequence 137, App
11	2146.5	86.9	475	US-10-080-334-138	Sequence 138, App
12	2139.5	86.7	475	US-10-225-810-41	Sequence 41, Appli
13	2137.5	86.6	475	US-10-225-810-18	Sequence 18, Appli
14	1914	77.5	422	US-10-291-737-4	Sequence 4, Appli
15	1914	77.5	422	US-10-365-564-4	Sequence 4, Appli
16	1722.5	69.8	483	US-10-225-810-19	Sequence 19, Appli
17	1718.5	69.6	483	US-10-392-219-2	Sequence 2, Appli
18	1718.5	69.6	507	US-10-467-685-7	Sequence 7, Appli
19	1696.5	68.7	478	US-10-225-810-14	Sequence 14, Appli
20	1677.5	67.9	481	US-10-225-810-22	Sequence 22, Appli
21	1476.5	59.8	427	US-10-225-810-20	Sequence 20, Appli
22	1448.5	58.7	427	US-10-291-737-2	Sequence 2, Appli
23	1448.5	58.7	427	US-10-365-564-2	Sequence 2, Appli
24	1416.5	57.4	477	US-10-225-810-17	Sequence 17, Appli
25	1263	51.2	263	US-10-080-334-24	Sequence 24, Appli
26	1255.5	50.9	485	US-10-055-025-2	Sequence 2, Appli
27	1255.5	50.9	485	US-10-154-419-92	Sequence 92, Appli

28	1255.5	50.9	485	6	US-11-043-889-52	Sequence 52, Appli
29	1255.5	50.9	504	4	US-10-080-334-139	Sequence 139, App
30	1252.5	50.7	476	4	US-10-168-651-23	Sequence 23, Appli
31	1252.5	50.7	476	4	US-10-755-889-22	Sequence 22, Appli
32	989.5	40.1	443	4	US-10-415-378-10	Sequence 10, Appli
33	972.5	39.4	369	4	US-10-108-260A-3609	Sequence 3609, Ap
34	866	35.1	482	4	US-10-225-810-34	Sequence 34, Appli
35	866	35.1	482	6	US-11-097-143-26835	Sequence 26835, A
36	866	35.1	504	4	US-10-080-334-141	Sequence 141, App
37	861	34.9	404	3	US-09-805-456-4	Sequence 4, Appli
38	861	34.9	404	3	US-09-805-456-5	Sequence 5, Appli
39	861	34.9	404	4	US-10-679-362-4	Sequence 4, Appli
40	861	34.9	404	4	US-10-679-362-5	Sequence 5, Appli
41	861	34.9	483	6	US-11-097-143-26193	Sequence 26193, A
42	861	34.9	486	6	US-11-097-143-31110	Sequence 31110, A
43	857	34.7	165	4	US-10-080-334-140	Sequence 140, App
44	853	34.5	404	4	US-10-291-737-5	Sequence 5, Appli
45	853	34.5	404	4	US-10-365-564-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1

US-09-805-456-2

; Sequence 2, Application US/09805456
; Publication No. US20030170778A1

; GENERAL INFORMATION:

; APPLICANT: WEI, Ming-Hui

; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,

; FILE OF INVENTION: AND USES THEREOF

; FILE REFERENCE: CL001062

; CURRENT APPLICATION NUMBER: US/09/805,456

; CURRENT FILING DATE: 2001-03-14

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 476

; TYPE: PRT

; ORGANISM: Human

US-09-805-456-2

Query Match 100.0%; Score 2469; DB 3; Length 476;
Best Local Similarity 100.0%; Pred. No. 1.7e-217;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MSTQRLRNEDYHYSSTDVSPEESPEGLNNLSPGSGYQRFQGSNSTTWFQTLHLKGN	60
Db	1	MSTQRLRNEDYHYSSTDVSPEESPEGLNNLSPGSGYQRFQGSNSTTWFQTLHLKGN	60
Qy	61	IGTGLGLPLAVKAGIVMGPIISLLIIGIVAVHCGILVKAHFCRLNKSFVDYGDIV	120
Db	61	IGTGLGLPLAVKAGIVMGPIISLLIIGIVAVHCGILVKAHFCRLNKSFVDYGDIV	120
Qy	121	MYGLESPCSWLRNHAHMGREVDVDFLIIVTQGLFCVYFVLADNFKQVIEAANGTTNNC	180
Db	121	MYGLESPCSWLRNHAHMGREVDVDFLIIVTQGLFCVYFVLADNFKQVIEAANGTTNNC	180
Qy	181	HNNETVILTPMTDRLYMLSPFLPVLVLFVTRNLRLALSIFSLLANITMLVSLVMIYQFIV	240
Db	181	HNNETVILTPMTDRLYMLSPFLPVLVLFVTRNLRLALSIFSLLANITMLVSLVMIYQFIV	240
Qy	241	QRIDPSPHLPLAVPKTYPLFEGTAIFSEFEGIGWVLPLENKMOKPRKPEPLILYLGWIVT	300
Db	241	QRIDPSPHLPLAVPKTYPLFEGTAIFSEFEGIGWVLPLENKMOKPRKPEPLILYLGWIVT	300
Qy	301	ILYISLGCGLGYLQFGANIQGSITLNLPCWLYQSVKLLYSIGIFFTYALQFVPAEIIIP	360
Db	301	ILYISLGCGLGYLQFGANIQGSITLNLPCWLYQSVKLLYSIGIFFTYALQFVPAEIIIP	360
Qy	361	FFVSRAPHELCBLVDLFRVTVLVCITLAILPRDLVIVSLVGSSSALALIIIPPLLE	420
Db	361	FFVSRAPHELCBLVDLFRVTVLVCITLAILPRDLVIVSLVGSSSALALIIIPPLLE	420

; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; FILE OF INVENTION: Using the Same
; CURRENT APPLICATION NUMBER: US/10/080,334
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/270,523
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/322,712
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 60/311,980
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/330,307
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/278,796
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: 60/281,521
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/276,677
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/311,595
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/270,220
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/274,295
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/318,526
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/286,548
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/291,765
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/270,797
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/276,400
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/270,810
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 22
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-080-334-22

Query Match 99.6%; Score 2460; DB 4; Length 496;
Best Local Similarity 99.8%; Pred. No. 1.2e-216;
Matches 475; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSTRQRNEDYHDSSTDVSPSESGNLNLSPPSGYQRFQSNSTTWFTQLIHLKGN 60
Db 21 MSTRQRNEDYHDSSTDVSPSESGNLNLSPPSGYQRFQSNSTTWFTQLIHLKGN 80

Qy 61 GTGGLGLPLAVKAGIWMGPISLLIIGIVAVCHMGILVKCAHFCRLNKSFDYDGTV 120
Db 81 GTGGLGLPLAVKAGIWMGPISLLIIGIVAVCHMGILVKCAHFCRLNKSFDYDGTV 140

Qy 121 MYGLESPCSWLRNHAHGRVDFFLIVTQLGFCVVYFVLADNFKQVIEAANGTTNCC 180
Db 141 MYGLESPCSWLRNHAHGRVDFFLIVTQLGFCVVYFVLADNFKQVIEAANGTTNCC 200

Qy 181 HNNETVILTPMDSRLYMLSPFLVLLVFNRLNLSIFSLANITMLVSLVMIYQFIV 240
Db 201 HNNETVILTPMDSRLYMLSPFLVLLVFNRLNLSIFSLANITMLVSLVMIYQFIV 260

Qy 241 QRIPDPHSLPLVAPWKTYPLFFGTAFSFEIGMWLPLLENKMKDPRKPFILYLGMWIVT 300
Db 261 QRIPDPHSLPLVAPWKTYPLFFGTAFSFEIGMWLPLLENKMKDPRKPFILYLGMWIVT 320

Qy 301 ILYISLGLGLYLFQGANIGSITLNLPCWLYQSCKLYSGIIFTYALQFVPAEIIIP 360
Db 321 ILYISLGLGLYLFQGANIGSITLNLPCWLYQSCKLYSGIIFTYALQFVPAEIIIP 380

Qy 361 FFVSRAPEHCELVVDLFRVTVLVCLTCLAILIPLRLDLVLSVSSSALALIIPPLE 420
Db 381 FFVSRAPEHCELVVDLFRVTVLVCLTCLAILIPLRLDLVLSVSSSALALIIPPLE 440

Qy 421 VTTFYSEGMSPLTIFKDALISILGFGVGVCTYEALEYLIOPSNAPIINSTCAFI 476
Db 441 VTTFYSEGMSPLTIFKDALISILGFGVGVCTYEALEYLIOPSNAPIINSTCAFI 496

RESULT 5
US-10-332-447-8
; Sequence 8, Application US/10332447
; Publication No. US20040053258A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; RAUMANN, Brigitte E.;
; APPLICANT: THORNTON, Michael; DING, Li; YUE, Henry;
; APPLICANT: TANG, Y. Tom; HARLAND, Lee; BURFORD, Neil;
; APPLICANT: GREENE, Barrie D.; SANJANWALA, Madhu S.;
; APPLICANT: BAUGHN, Mariah R.; YAO, Monique G.; YANG, Junming;
; APPLICANT: ARVIZU, Chandra S.; GANDHI, Ameena R.;
; APPLICANT: HAFALIA, April J.A.; TRIBOULEY, Catherine M.;
; APPLICANT: WALSH, Roderick T.; RAMKUMAR, Javalaxmi;
; APPLICANT: LU, Yan; LU, Dyung Aina M.; AZIMZAI, Yalda;
; APPLICANT: LAL, Preeti; ELLIOTT, Vicki S.; NGUYEN, Dannel B.;
; APPLICANT: XU, Yuming; SEILHAMER, Jeffrey J.; BOROWSKY, Mark L.;
; APPLICANT: KHAN, Farrah A.; KEARNEY, Liam; THANGAVELU, Kavitha;
; APPLICANT: DAS, Debopriya; POLICKY, Jennifer L.
; TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS
; FILE REFERENCE: PI-0149 USN
; CURRENT APPLICATION NUMBER: US/10/332,447
; CURRENT FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: US 60/216,547
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/218,232
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/220,112
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: US 60/221,839
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PERL Program
; SEQ ID NO 8
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID NO. US20040053258A1 7476747CD1
US-10-332-447-8

Query Match 97.3%; Score 2402.5; DB 4; Length 568;
Best Local Similarity 88.3%; Pred. No. 2.6e-211;
Matches 474; Conservative 0; Mismatches 2; Indels 61; Gaps 2;

Qy 1 MSTRQRNEDYHDSSTDVSPSESGNLNLSPPSGYQRFQSNSTTWFTQLIHLKGN 60
Db 32 MSTRQRNEDYHDSSTDVSPSESGNLNLSPPSGYQRFQSNSTTWFTQLIHLKGN 91

Qy 61 GTGGLGLPLAVKAGIWMGPISLLIIGIVAVCHMGILVKCAHFCRLNKSFDYDGTV 120
Db 92 GTGGLGLPLAVKAGIWMGPISLLIIGIVAVCHMGILVKCAHFCRLNKSFDYDGTV 151

Qy 121 MYGLESPCSWLRNHAHGRVDFFLIVTQLGFCVVYFVLADNFKQVIEAANGTTNCC 180
Db 152 MYGLESPCSWLRNHAHGRVDFFLIVTQLGFCVVYFVLADNFKQVIEAANGTTNCC 211

Qy 181 HNNETVILTPMDSRLYMLSPFLVLLVFNRLNLSIFSLANITMLVSLVMIYQFIV 240
Db 212 HNNETVILTPMDSRLYMLSPFLVLLVFNRLNLSIFSLANITMLVSLVMIYQFIV 271

Qy 241 QRIPDPHSLPLVAPWKTYPLFFGTAFSFEIGMWLPLLENKMKDPRKPFILYLGMWIVT 300


```
; SEQ ID NO 7
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Human
US-09-805-456-7

Query Match      94.1%; Score 2324; DB 3; Length 476;
Best Local Similarity 94.5%; Pred. No. 3.3e-204;
Matches 450; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

QY 1 MSTQRLNEDYHDYSSSTDVSPSESGEGLNLSPPGYSQRFQGSNSTTWFQTLIHLKGN 60
DB 1 MSTQRLNEDYHDYSSSTDVSPSESGEGLNLSPPGYSQRFQGSNSTTWFQTLIHLKGN 60

QY 61 IGTGLLGLPLAVKNAAGIVMGPIISLLIIGIVAVHCHMGILVKCAHFCRRLNKSFVDYGDV 120
DB 61 IGTGLLGLPLAVKNAAGIVMGPIISLLIIGIVAVHCHMGILVKCAHFCRRLNKSFVDYGDV 120

QY 121 MYGLESPCSWLNRHAHWGRVDFFLIVTQLGFCVYFVLADNFKQVIEAANGTTNNC 180
DB 121 MYGLESPCSWLNRHAHWGRVDFFLIVTQLGFCVYFVLADNFKQVIEAANGTTNNC 180

QY 181 HNETVILTPMDSRLYMLSPFLPVLVLFVTRNLRALSIFSLLANITMLVSLVMIYQFIV 240
DB 181 HNETVILTPMDSRLYMLSPFLPVLVLFVTRNLRALSIFSLLANITMLVSLVMIYQFIV 240

QY 241 QRIIDPDPHPLVAPWKTYPLFFGTAFISFEGIGMWLPLENKMKDPRKPLILYLGMVIVT 300
DB 241 QRIIDPDPHPLVAPWKTYPLFFGTAFISFEGIGMWLPLENKMKDPRKPLILYLGMVIVT 300

QY 301 ILYISLGLCYLQFGANIQGSITLNLPCNLVYQSVKLLYSIGIFTYALQFVVPAAIIP 360
DB 301 ILYISLGLCYLQFGANIQGSITLNLPCNLVYQSVKLLYSIGIFTYALQFVVPAAIIP 360

QY 361 FFVSRAPBCHBLVDLFRVTVLVCITCLAILIPRLDLVILVSGSVSSALALIIPPLLE 420
DB 361 VIVSWCKCTLMVDLGGISAMLCCKTCLAILIPRLDLVILVSGSVSSALALIIPPLLE 420

QY 421 VTTFYSEGMSPLTIFKDALISILGVGFVGVGTYEALYELIOPSNAPIFINSTCAFI 476
DB 421 VTTFYSEGMSPLTIFKDALISILGVGFVGVGTYEALYELIOPSNAPIFINSTCAFI 476

RESULT 8
US-10-679-362-7
; Sequence 7, Application US/10679362
; Publication No. US20040146887A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS.
; FILE REFERENCE: CLO01062CON
; CURRENT APPLICATION NUMBER: US/10/679,362
; CURRENT FILING DATE: 2003-10-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Human
US-10-679-362-7

Query Match      94.1%; Score 2324; DB 4; Length 476;
Best Local Similarity 94.5%; Pred. No. 3.3e-204;
Matches 450; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

QY 1 MSTQRLNEDYHDYSSSTDVSPSESGEGLNLSPPGYSQRFQGSNSTTWFQTLIHLKGN 60
DB 1 MSTQRLNEDYHDYSSSTDVSPSESGEGLNLSPPGYSQRFQGSNSTTWFQTLIHLKGN 60

QY 61 IGTGLLGLPLAVKNAAGIVMGPIISLLIIGIVAVHCHMGILVKCAHFCRRLNKSFVDYGDV 120
DB 61 IGTGLLGLPLAVKNAAGIVMGPIISLLIIGIVAVHCHMGILVKCAHFCRRLNKSFVDYGDV 120

QY 121 MYGLESPCSWLNRHAHWGRVDFFLIVTQLGFCVYFVLADNFKQVIEAANGTTNNC 180
DB 121 MYGLESPCSWLNRHAHWGRVDFFLIVTQLGFCVYFVLADNFKQVIEAANGTTNNC 180

QY 181 HNETVILTPMDSRLYMLSPFLPVLVLFVTRNLRALSIFSLLANITMLVSLVMIYQFIV 240
DB 181 HNETVILTPMDSRLYMLSPFLPVLVLFVTRNLRALSIFSLLANITMLVSLVMIYQFIV 240

QY 241 QRIIDPDPHPLVAPWKTYPLFFGTAFISFEGIGMWLPLENKMKDPRKPLILYLGMVIVT 300
DB 241 QRIIDPDPHPLVAPWKTYPLFFGTAFISFEGIGMWLPLENKMKDPRKPLILYLGMVIVT 300

QY 301 ILYISLGLCYLQFGANIQGSITLNLPCNLVYQSVKLLYSIGIFTYALQFVVPAAIIP 360
DB 301 ILYISLGLCYLQFGANIQGSITLNLPCNLVYQSVKLLYSIGIFTYALQFVVPAAIIP 360

QY 361 FFVSRAPBCHBLVDLFRVTVLVCITCLAILIPRLDLVILVSGSVSSALALIIPPLLE 420
DB 361 VIVSWCKCTLMVDLGGISAMLCCKTCLAILIPRLDLVILVSGSVSSALALIIPPLLE 420

QY 421 VTTFYSEGMSPLTIFKDALISILGVGFVGVGTYEALYELIOPSNAPIFINSTCAFI 476
DB 421 VTTFYSEGMSPLTIFKDALISILGVGFVGVGTYEALYELIOPSNAPIFINSTCAFI 476

US-10-094-749-3076
; Sequence 3076, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOKYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3076
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-3076

Query Match      89.8%; Score 2217; DB 4; Length 434;
Best Local Similarity 90.8%; Pred. No. 1.9e-194;
Matches 432; Conservative 1; Mismatches 1; Indels 42; Gaps 1;

QY 1 MSTQRLNEDYHDYSSSTDVSPSESGEGLNLSPPGYSQRFQGSNSTTWFQTLIHLKGN 60
DB 1 MSTQRLNEDYHDYSSSTDVSPSESGEGLNLSPPGYSQRFQGSNSTTWFQTLIHLKGN 60
```


APPLICANT: Li, Li
 APPLICANT: Shenoy, Suresh G
 APPLICANT: Kekuda, Ramesh
 APPLICANT: Spytek, Kimberly A
 APPLICANT: Vernet, Corine A. M.
 APPLICANT: Malyankar, Uriel M
 APPLICANT: Guo, Xiaojia
 APPLICANT: Gusev, Vladimir Y
 APPLICANT: Casman, Stacie J
 APPLICANT: Boldog, Ferenc L
 APPLICANT: Furtak, Katarzyna
 APPLICANT: Tchernev, Velizar T
 APPLICANT: Patturajan, Meera
 APPLICANT: Padigaru, Muralidhara
 APPLICANT: Gangoli, Esha A
 APPLICANT: Liu, Xiaohong
 APPLICANT: Baumgartner, Jason C.
 APPLICANT: Gerlach, Valerie
 APPLICANT: Spaderna, Steven K
 APPLICANT: Zernusen, Bryan D
 TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
 TITLE OF INVENTION: Using the Same
 FILE REFERENCE: 21402-275
 CURRENT APPLICATION NUMBER: US/10/080,334
 CURRENT FILING DATE: 2002-02-21
 PRIOR APPLICATION NUMBER: 60/270,523
 PRIOR FILING DATE: 2001-02-21
 PRIOR APPLICATION NUMBER: 60/322,712
 PRIOR FILING DATE: 2001-09-17
 PRIOR APPLICATION NUMBER: 60/311,980
 PRIOR FILING DATE: 2001-08-13
 PRIOR APPLICATION NUMBER: 60/330,307
 PRIOR FILING DATE: 2001-10-18
 PRIOR APPLICATION NUMBER: 60/278,796
 PRIOR FILING DATE: 2001-03-26
 PRIOR APPLICATION NUMBER: 60/281,521
 PRIOR FILING DATE: 2001-04-04
 PRIOR APPLICATION NUMBER: 60/276,677
 PRIOR FILING DATE: 2001-03-16
 PRIOR APPLICATION NUMBER: 60/311,595
 PRIOR FILING DATE: 2001-08-10
 PRIOR APPLICATION NUMBER: 60/270,220
 PRIOR FILING DATE: 2001-02-21
 PRIOR APPLICATION NUMBER: 60/274,295
 PRIOR FILING DATE: 2001-03-08
 PRIOR APPLICATION NUMBER: 60/318,526
 PRIOR FILING DATE: 2001-09-10
 PRIOR APPLICATION NUMBER: 60/286,548
 PRIOR FILING DATE: 2001-04-25
 PRIOR APPLICATION NUMBER: 60/291,765
 PRIOR FILING DATE: 2001-05-17
 PRIOR APPLICATION NUMBER: 60/270,797
 PRIOR FILING DATE: 2001-02-23
 PRIOR APPLICATION NUMBER: 60/276,400
 PRIOR FILING DATE: 2001-03-16
 PRIOR APPLICATION NUMBER: 60/270,810
 PRIOR FILING DATE: 2001-02-23
 NUMBER OF SEQ ID NOS: 388
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 138
 LENGTH: 475
 TYPE: PRT
 ORGANISM: Rattus norvegicus
 US-10-080-334-138
 Query Match 86.9%; Score 2146.5; DB 4; Length 475;
 Best Local Similarity 85.7%; Pred. No. 6.3e-188;
 Matches 408; Conservative 36; Mismatches 31; Indels 1; Gaps 1;
 1 MSTQRLRNEDYHDYSTDVSPESPSEGLNNLSPGSGYQRFQGSNSTTWFOTLIHLKGN 60
 1 MSTQRLRNEDYHDYSTDVSPESPSEGLGSGF-SPGSYQRLGSENSMTWQTLIHLKGN 59
 APPLICANT: Li, Li
 APPLICANT: Shenoy, Suresh G
 APPLICANT: Kekuda, Ramesh
 APPLICANT: Spytek, Kimberly A
 APPLICANT: Vernet, Corine A. M.
 APPLICANT: Malyankar, Uriel M
 APPLICANT: Guo, Xiaojia
 APPLICANT: Gusev, Vladimir Y
 APPLICANT: Casman, Stacie J
 APPLICANT: Boldog, Ferenc L
 APPLICANT: Furtak, Katarzyna
 APPLICANT: Tchernev, Velizar T
 APPLICANT: Patturajan, Meera
 APPLICANT: Padigaru, Muralidhara
 APPLICANT: Gangoli, Esha A
 APPLICANT: Liu, Xiaohong
 APPLICANT: Baumgartner, Jason C.
 APPLICANT: Gerlach, Valerie
 APPLICANT: Spaderna, Steven K
 APPLICANT: Zernusen, Bryan D
 TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
 TITLE OF INVENTION: Using the Same
 FILE REFERENCE: 21402-275
 CURRENT APPLICATION NUMBER: US/10/080,334
 CURRENT FILING DATE: 2002-02-21
 PRIOR APPLICATION NUMBER: 60/270,523
 PRIOR FILING DATE: 2001-02-21
 PRIOR APPLICATION NUMBER: 60/322,712
 PRIOR FILING DATE: 2001-09-17
 PRIOR APPLICATION NUMBER: 60/311,980
 PRIOR FILING DATE: 2001-08-13
 PRIOR APPLICATION NUMBER: 60/330,307
 PRIOR FILING DATE: 2001-10-18
 PRIOR APPLICATION NUMBER: 60/278,796
 PRIOR FILING DATE: 2001-03-26
 PRIOR APPLICATION NUMBER: 60/281,521
 PRIOR FILING DATE: 2001-04-04
 PRIOR APPLICATION NUMBER: 60/276,677
 PRIOR FILING DATE: 2001-03-16
 PRIOR APPLICATION NUMBER: 60/311,595
 PRIOR FILING DATE: 2001-08-10
 PRIOR APPLICATION NUMBER: 60/270,220
 PRIOR FILING DATE: 2001-02-21
 PRIOR APPLICATION NUMBER: 60/274,295
 PRIOR FILING DATE: 2001-03-08
 PRIOR APPLICATION NUMBER: 60/318,526
 PRIOR FILING DATE: 2001-09-10
 PRIOR APPLICATION NUMBER: 60/286,548
 PRIOR FILING DATE: 2001-04-25
 PRIOR APPLICATION NUMBER: 60/291,765
 PRIOR FILING DATE: 2001-05-17
 PRIOR APPLICATION NUMBER: 60/270,797
 PRIOR FILING DATE: 2001-02-23
 PRIOR APPLICATION NUMBER: 60/276,400
 PRIOR FILING DATE: 2001-03-16
 PRIOR APPLICATION NUMBER: 60/270,810
 PRIOR FILING DATE: 2001-02-23
 NUMBER OF SEQ ID NOS: 388
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 138
 LENGTH: 475
 TYPE: PRT
 ORGANISM: Rattus norvegicus
 US-10-080-334-138
 Query Match 86.9%; Score 2146.5; DB 4; Length 475;
 Best Local Similarity 85.7%; Pred. No. 6.3e-188;
 Matches 408; Conservative 36; Mismatches 31; Indels 1; Gaps 1;
 1 MSTQRLRNEDYHDYSTDVSPESPSEGLNNLSPGSGYQRFQGSNSTTWFOTLIHLKGN 60
 1 MSTQRLRNEDYHDYSTDVSPESPSEGLGSGF-SPGSYQRLGSENSMTWQTLIHLKGN 59

Db 300 VLYISLSGLYQPGADIKGSITLNLPCWLYQSVKLLYSIGIFFYFVALQFYVAEEIIP 359
Qy 361 FFVSRAPHELCVVDLFRVTVLVCLTCLAILPRDLVLSLGVSSSSALALIIPPLE 420
Db 360 AIVSRVPEFELVVDLSARTAMVCTCVLAVLPRDLVLSLGVSSSSALALIIPPLE 419
Qy 421 VTFYSGMSPLTIFKDALISILGFVGVGTVEALYELIQPSNAPIFINSTAFI 476
Db 420 VTTYGEGISPLTTTKDALISILGFVGVGTVEALYELIQPSNAPIFINSTAFI 475

RESULT 13
US-10-225-810-18
; Sequence 18, Application US/10225810
; Publication No. US20030157512A1
; GENERAL INFORMATION:
; APPLICANT: Bermingham, Jr., John R.
; TITLE OF INVENTION: Tramdorins and Methods of Using Tramdorin
; FILE REFERENCE: McLaugh-07165
; CURRENT APPLICATION NUMBER: US/10/225,810
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-225-810-18

Query Match 86.6%; Score 2137.5; DB 4; Length 475;
Best Local Similarity 85.9%; Pred. No. 4.2e-187;
Matches 409; Conservative 32; Mismatches 34; Indels 1; Gaps 1;

Qy 1 MSTORLNEDYHDSSTDVSPERSPSRGLNLSPPGSQVRFQGSNTWFTQTLHLKGN 60
Db 1 MSTORLNEDYHDSSTDVSPERSPSRGLGSP-SPGSYQRLGESSNTWFTQTLHLKGN 59

Qy 61 IGTGLLGLPLAVKNAGIVMGPIISLLIIGIVAVHCMGILVKCAHHFCRLNKSFDYDGTV 120
Db 60 IGTGLLGLPLAVKNAGLLGLPLSLVIGIVAVHCMGILVKCAHHLCRLNKPFLDYDGTV 119

Qy 121 MYGLESPCSWLNRHAGRRVDFLLIVTQLGFCVCYFVFLADNFKQVIEAANGTTNCC 180
Db 120 MYGLESPSTWVRNHSWGRRIIVDFLLIVTQLGFCVCYFVFLADNFKQVIEAANGTTTNC 179

Qy 181 HNNETVLTPTMDSRLYMLGFLPLVLLVFTNRNLALSIFSLLANITMLVSLVMYQFIV 240
Db 180 NNNVTVIPTMDSRLYMPFLPLVLLVFTNRNLALSIFSLLANISMFVSLMIYQFIV 239

Qy 241 QRIPDPShLPLVAPWKTYPLFFGTAIFSGEIGMVLPLENKMKDPRKPLILYLGWVIT 300
Db 240 QRIPDPShLPLVAPWKTYPLFFGTAIFAFEGIGVVLPLENKMKDSQKFPLILYLGMAIT 299

Qy 301 ILYISLGLGLYQPGANIQSITLNLPCWLYQSVKLLYSIGIFFYFVALQFYVAEEIIP 360
Db 300 VLYISLSGLYQPGANIKGSITLNLPCWLYQSVKLLYSIGIFFYFVALQFYVAEEIIP 359

Qy 361 FFVSRAPHELCVVDLFRVTVLVCLTCLAILPRDLVLSLGVSSSSALALIIPPLE 420
Db 360 AIVSRVPEFELVVDLCVTRAMVCTCVLAILPRDLVLSLGVSSSSALALIIPPLE 419

Qy 421 VTFYSGMSPLTIFKDALISILGFVGVGTVEALYELIQPSNAPIFINSTAFI 476
Db 420 VTTYGEGISPLTTTKDALISILGFVGVGTVEALYELIQPSNAPIFINSTAFI 475

RESULT 14
US-10-291-737-4
; Sequence 4, Application US/10291737
; Publication No. US20030087299A1
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001017-CON
; CURRENT APPLICATION NUMBER: US/10/291,737
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 09/822,859
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: US 60/254,588
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Human
US-10-291-737-4

Query Match 77.5%; Score 1914; DB 4; Length 422;
Best Local Similarity 86.0%; Pred. No. 1.1e-166;
Matches 363; Conservative 33; Mismatches 26; Indels 0; Gaps 0;

Qy 51 QTLIHLKGNIGTGLLGLPLAVKNAGIVMGPIISLLIIGIVAVHCMGILVKCAHHFCRLN 110
Db 1 QTLIHLKGNIGTGLLGLPLAVKNAGLLGLPLSLVIGIVAVHCMGILVKCAHHLCRLN 60

Qy 111 KSPVDYDGTVMYGLSPCSWLNRHAGRRVDFLLIVTQLGFCVCYFVFLADNFKQVI 170
Db 61 KPFLDYDGTVMYGLSPSTWIRNHSWGRRIIVDFLLIVTQLGFCVCYFVFLADNFKQVI 120

Qy 171 EAANGTTNCHNETVLTPTMDSRLYMLGFLPLVLLVFTNRNLALSIFSLLANITMLV 230
Db 121 EAANGTTNCHNETVLTPTMDSRLYMLTFLPLVLLSFIRNLRLISIFSLLANISMFV 180

Qy 231 SLVMYQFIVQRIQIPDPShLPLVAPWKTYPLFFGTAIFSGEIGMVLPLENKMKDPRKPL 290
Db 181 SLIMIYQFIVQRIQIPDPShLPLVAPWKTYPLFFGTAIFAFEGIGVVLPLENKMKDSQKFPL 240

Qy 291 ILYGMVITVILYISLGLYQPGANIQSITLNLPCWLYQSVKLLYSIGIFFYFVALQ 350
Db 241 ILYGMAITVILYISLGLYQPGADIKGSITLNLPCWLYQSVKLLYSIGIFFYFVALQ 300

Qy 351 FYVPAEIIIPFFVSRAPHELCVVDLFRVTVLVCLTCLAILPRDLVLSLGVSSSSA 410
Db 301 FYVAAEIIIPAVISRVPERFELVVDLSARTAMVCTCVLAVLPRDLVLSLGVSSSSA 360

Qy 411 LALIIPPLEVITTFYSGMSPLTIFKDALISILGFVGVGTVEALYELIQPSNAPIFIN 470
Db 361 LALIIPPLKVTYYGEGISPLTITKDALISILGFVGVGTVEALYELIQPSHSDSSTN 420

Qy 471 ST 472
Db 421 ST 422

RESULT 15
US-10-365-564-4
; Sequence 4, Application US/10365564
; Publication No. US20030143623A1
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; FILE REFERENCE: AND USES THEREOF
; FILE REFERENCE: CL001017-CIPCON
; CURRENT APPLICATION NUMBER: US/10/365,564
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: US 09/822,859
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: US 60/254,588
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4

LENGTH: 422
TYPE: PRT
ORGANISM: Human
US-10-365-564-4

Query Match 77.5%; Score 1914; DB 4; Length 422;
Best Local Similarity 86.0%; Pred. No. 1.1e-166;
Matches 363; Conservative 33; Mismatches 26; Indels 0; Gaps 0;

Qy 51 QTLHLKGNIGTGLGPLAVKAGIYMGPIISLLIIGIVAVHCGIILVKCAHFCRELN 110
Db 1 QTLHLKGNIGTGLGPLAVKAGLGLGSLSLVIGIVAVHCGIILVKCAHLCRELN 60
Qy 111 KSFVDYDGTVMYGLESSPCSWLRNHAHWGRVDFFLIVTQLGFCVYFVFLADNFKQVI 170
Db 61 KPFLDYDGTVMYGLECSFSTWIRNHSWGRRIVDFFLVVTQLGFCVYFVFLADNFKQVI 120
Qy 171 EAAAGTTNCHNNETVILTPMDSRLYMLSPFLPFLVLLVFIRNLRALSIFSLANITMLV 230
Db 121 EAAAGTTNCHNNETVILTPMDSRLYMLTFLPFLVLLSFIRNLRILSFSLANISMFV 180
Qy 231 SLVMYQRIVORIPDPSHPLVAPWKTYPLFEGTAFISFEGIGMVLPLENKKMDRKPPL 290
Db 181 SLMIYQRIVORIPDPSHPLVAPWKTYPLFEGTAFISFEGIGMVLPLENKKMDRKPPL 240
Qy 291 ILYLGMVITVILYISLGLGYLQFGANTQGSITLNLPCWLYQSVKLLYSIGIFTYALQ 350
Db 241 ILYLGMVITVILYISLGLGYLQFGADIKGSITLNLPCWLYQSVKLLYSIGIFTYALQ 300
Qy 351 FYVPAEIIIPFVSRAPEHCELVDL FVRTVLVCLTCLAILIPRLDLVISLVGSVSSA 410
Db 301 FYVAAEIIIPAVSRVPERFELVDLSARTAMVCVLCVAVLIPRLDLVISLVGSVSSA 360
Qy 411 LALIIPPLEVTTFYSEGMSPITPKDALISILGFVGVGTVEALYELIOPSNAPIETN 470
Db 361 LALIIPPLEKVTYYEGISPLTITKDALISILGFVGVGTVEALYELIOPSNAPIETN 420
Qy 471 ST 472
Db 421 ST 422

Search completed: August 24, 2006, 01:29:45
Job time : 186 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2006, 01:26:49 ; Search time 33 Seconds
(without alignments)
977.303 Million cell updates/sec

Title: US-10-679-362-2
Perfect score: 2469
Sequence: 1 MSTQRLRNEDYHDSYTDVS.....YELIQSNAPIFINSTCAFI 476

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 243793 seqs, 67754213 residues

Total number of hits satisfying chosen parameters: 243793

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New.*

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_NEW_PUB.pap.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US06_NEW_PUB.pap.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_NEW_PUB.pap.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_NEW_PUB.pap.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/PCT_NEW_PUB.pap.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10_NEW_PUB.pap.*
- 7: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_NEW_PUB.pap.*
- 8: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US60_NEW_PUB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1255.5	50.9	504	6	US-10-505-928-42
2	972.5	39.4	369	7	US-11-056-355B-75678
3	465.5	18.9	436	6	US-10-953-349-368
4	410.5	16.6	371	6	US-10-953-349-369
5	405	16.4	425	6	US-10-449-902-37011
6	405	16.4	425	6	US-10-449-902-49196
7	404.5	16.4	422	6	US-10-449-902-32842
8	404.5	16.4	422	6	US-10-449-902-53550
9	402.5	16.3	422	6	US-10-449-902-53486
10	400	16.2	395	6	US-10-449-902-29254
11	382	15.5	355	6	US-10-953-349-370
12	339.5	13.8	342	6	US-10-449-902-32686
13	332	13.4	349	6	US-10-449-902-36917
14	318	12.9	531	7	US-11-174-307B-5170
15	302	12.2	468	7	US-11-056-355B-105677
16	302	12.2	468	7	US-11-056-355B-116916
17	302	12.2	474	7	US-11-056-355B-105676
18	302	12.2	474	7	US-11-056-355B-116915
19	302	12.2	550	7	US-11-056-355B-105675
20	302	12.2	550	7	US-11-056-355B-116914
21	300.5	12.2	436	7	US-11-056-355B-54601
22	300.5	12.2	449	7	US-11-056-355B-54600
23	300.5	12.2	458	7	US-11-056-355B-54599
24	292	11.8	471	7	US-11-056-355B-79663
25	292	11.8	526	7	US-11-056-355B-79662

RESULT 1

US-10-505-928-42
; Sequence 42, Application US/10505928
; Publication No. US20060088532A1

GENERAL INFORMATION:

; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 42
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-505-928-42

Query Match 50.9%; Score 1255.5; DB 6; Length 504;
Best Local Similarity 52.0%; Pred. No. 5.4e-96;
Matches 254; Conservative 80; Mismatches 119; Indels 35; Gaps 6;

Qy	6	LRNEDYHDSYTDVSPPEESPESEGLNNLSPEGSYQRFQGSNTTWFQTLIHLKKNIGTGL	65
Db	26	LINEQNFDTGSDERHEQE-----LLPVQKHVQLDDQEGISFVQTLMLHLKKNIGTGL	77
Qy	66	LGLPLAVNAGIVGMPISLLIIGIVAVCMGILVKCAHFCRRLNKSFVDYDGTVMYCLE	125
Db	78	LGLPLAKNAGIVLGPISLVFIGIISVHCMLVRCRSHFLCLRFKFKSTGLSYDTSVFAME	137
Qy	126	SSPCSRLNHAHGRVVDVFLVITQLGFCVCYVFELADNFKOV-----IEAANG	175
Db	138	VSPWSCIQKQAAMGRSVVDVFLVITQLGFCVSVIVFLAENVKQVHGFLESKVFISNSTN	197
Qy	176	TTNNCHNNTVILPTTMDSLRYMLSLPFLVLLVFIIRNLRLALSIFSLANITMLVSLYMI	235
Db	198	SSNFCERR-----SVDLRIYMLCFPLPFIALLVFIRELKNLFLVLSFLANYSMAVSLVII	250
Qy	236	YQFTVQRPDPSPHPLVAPWKTYPLFTGTATFSEGIQGVLPLENKMDPKRFPFLIYVLG	295
Db	251	YQYVVRNMPDHPNLPVAGWKYPLFTGTAVFAFEGIGVLPLENQMKESKRFPQALNIG	310
Qy	296	MVITVILYISLGCILQYQFQGNIOGSITNLPL-NCWLYQSVKLLYSIGIFFTYALQFVVP	354
Db	311	MGIVTTLYTLATLGYMCFHDEIKGISTNLPLQDVMVLYQSVKILYSIGIFVTSIQFVVP	370

ALIGNMENTS

Sequence 75679, A
Sequence 75678, A
Sequence 75677, A
Sequence 96651, A
Sequence 96650, A
Sequence 96649, A
Sequence 38517, A
Sequence 44879, A
Sequence 34888, A
Sequence 8582, Ap
Sequence 34887, A
Sequence 8581, Ap
Sequence 34886, A
Sequence 8580, Ap
Sequence 36518, A
Sequence 44880, A
Sequence 42650, A
Sequence 38581, A
Sequence 52309, A
Sequence 79664, A

QY 355 ABIIIPFVSRAPEHCELVVDLQVTVVCLTCLAILIPRLDLVSLGVSSSALALI 414
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 371 AEIIIPGITSKFHTKWKQICEFGIRSLFSITCAGAILIPRLDIVISFVGAVSSSTALI 430
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
QY 415 IPLLLEVTTFYSEGMSPLTIFFKDALISILGFVGVGTVEALYELIOPS-----NAP 466
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 431 LPPDLVELTTSKEHYNIMWLKNISIAFTGVGVFLGTITVEEIIYTPKVVAGTPQSP 490
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
QY 467 IF-INSTC 473
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 491 FLNLNSTC 498
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
RESULT 2
US-11-293-697-3609
; Sequence 3609, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3609
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-3609

Query Match 39.4%; Score 972.5; DB 7; Length 369;
Best Local Similarity 53.0%; Pred. No. 9e-73;
Matches 196; Conservative 63; Mismatches 84; Indels 27; Gaps 5;
QY 124 LESSPCSWLRNHAHWRRVVDFFLIIVTQGLGCCYVFFVLADNFKQV-----IEAA 173
: | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 1 MEVSPWCLQQAAGRSVVDFFLIIVTQGLGFCSVIIVFLAENKQVHEGFLESKVFISS 60
: | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
QY 174 NGTNNCHNETVILTMTDSRLYMLSLFPLVLVIRNLRLSIFSLLANITMLVSLV 233
: | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 61 TNSNPNCR-----SVDLRIYMLCLFPFIILLVFIRELKNLFLVLSFLANVSMVSLV 113
: | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
QY 234 MIYQFIVQRIIDPDSHLPLVAPWKTYPLFFGTAFISFEGIGMVLPLENKMKDPRKFPILY 293
: | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 114 IYQVVRNPDPNPLPIVAGKKYPLFFGTAFVAFEGIGVVLPLENQMKESKFFPQALN 173
: | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
QY 294 LGMVIVTILYSLGCLGYLQFGANIOGSITLNLPL-NCWLYQSVKLLYSIGIFFTYALQFY 352
: | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 174 IGMGIIVTLVTLATLQVCMFRDEIKGSITLNLQPDVWLYQSVKILYSFGIFVTSIQFY 233
: | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
QY 353 VPABIIIPFVSRAPEHCELVVDLQVTVVCLTCLAILIPRLDLVSLGVSSSALA 412
: | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 234 VPABIIIPGITSKFTKWKQICEFGIRSLFSITCAGAILIPRLDIVISFVGAVSSSTLA 293
: | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
QY 413 LIIPPLEVTTFYSEGMSPLTIFFKDALISILGFVGVGTVEALYELIOPS-----N 464
: | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 294 LILPPLVELTTSKEHYNIMWLKNISIAFTGVGVFLGTITVEEIIYTPKVVAGTPQ 353
: | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
QY 465 APIF-INSTC 473
: | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 354 SPFLNLNSTC 363
: | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :

RESULT 3
US-10-953-349-368
; Sequence 368, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES

; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 368
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-368

Query Match 18.9%; Score 465.5; DB 6; Length 436;
Best Local Similarity 32.4%; Pred. No. 8.2e-31;
Matches 143; Conservative 66; Mismatches 155; Indels 77; Gaps 14;
QY 50 FOTLIHLKGNIGTGLGLPLAVKNAGIVMGPSILTIIGIVAVHCGMILVKCAHFCRRRL 109
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 37 FKTFANVFIAVVGAGVGLPYAFKRTGWLGMVLLVSVSLTHHCMLLV-----YTRRK 91
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 110 NKSP-----VDYGDVTVMYGLESPSCSWLRNHAHWRRVVDFFLIIVTQGLGCCYVFFVL 162
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 92 LDSFNAGISKIGSGD-----LGFVCGSL-----GRIVVDLFIILSQAGFCVGLIFI 140
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 163 ADNFQVIEAANGTT--NNCHNETVILTPT-----MDSLYMLSLFPL 205
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 141 -----GTTLANLSDPE-----SPTSLRHQTRLGSEFLGVSSKSLYIWGCFPFQ 184
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 206 VLLVFIIRNLRLSIFSLLANI-----TMLVSLVMYQFIVQRIIDPDSHLPLVAPWKTYPLF 261
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 185 LGLNSIKTTHLAPLSIFADIVDLGAMAVVIVEDSMILKQRPD-----VFAFGMSLF 238
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 262 ---FGTAIFSEFEGIGMVLPLENKMKDPRKFPILYLGWIVTILYISLGLCYLQFGANI 318
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 239 LYGMGVAVYSFEGVMVLPLESEMCKDKFGKVLALGMGFIISLIYIAFGILGYLAFGEDT 298
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 319 QGSITLNLPCWLYQSVKLLYSIGIFFTYALQFYVPAEIIIPFVSRAPEHCELVVDLQV 378
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 299 MDIITANLQAGLVSTVQVGLGCINLFTPLMNNPVFEIVERRP--SRG-----MYSAWL 351
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 379 RTVLVCLTCLAILIPRLDLVSLGVSSSALALIIPPLEVTTFYSEGMSPLTIFFKDA 438
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 352 RWVLVLAIVLVALFVFNADFLSLVSSCCCVLGFVLPAFLHLLVF-KEEMGWLQWSSDT 410
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 439 LISILGFVGVGTVEALYEL 459
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 411 AIVVLGWLAVSGTWSSLSSEI 431
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

RESULT 4
US-10-953-349-369
; Sequence 369, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 369
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-369

Query Match 16.6%; Score 410.5; DB 6; Length 371;
Best Local Similarity 32.0%; Pred. No. 2.4e-26;
Matches 132; Conservative 60; Mismatches 143; Indels 77; Gaps 14;
QY 79 MGPSILLIIGIVAVHCGMILVKCAHFCRRRLNSF-----VDYGDVTVMYGLESPSCSW 131

```
Db 1 MGVLVSVSVLTHCMMLLV-----YTRRKLDSEFNAGISKIGSGD-----LGFVAVGCS 50
Qy 132 LRNHAHGRVVDVDFLIYVTLQGFCCVYFVFLADNFKQVIEAANGTT--NNCHNNETVILTP 190
Db 51 L-----GRIVVDLFIILSQAGFCVGLIFI-----GTTLANLSDPE-----SP 88
Qy 191 T-----MDSLYMLSLFPLFVLVLLVFINRLALSIFSLANI-----TMLV 230
Db 89 TSLRHQFTRLGSEFLGVSSKSLYIAGFCVGLIFI-----GTTLANLSDPE-----SP 148
Qy 231 SLVMIYQIVQRIIDPSPHPLVAPWKTYPILP---FGTAIFSEFGIGMVLPLENKKDKRK 287
Db 149 VIVEDSMILKQRPD-----VWAFGMSFLYGMGVAVYFEGVGMVLPLESEMKDKDK 202
Qy 288 PFLILYLGMIWITLYISLGLCLYQFGANIGSITLNLNPNCLWYQSVKLLYSIGIFTY 347
Db 203 PGKVALGGMGFSIYIYAGFGLVLAAGEDTMDITANLGLAGLYSTVVQLGCLINLFTTF 262
Qy 348 ALQFYVPAEIIIPFVSRAPEHCELVVDLFRVTVLVCITCLAILIPRLDLVISLVGSVS 407
Db 263 PLMNPVFEIVERRF-SRG-----MYSAMLRWVLVLAVALFVFPNFADFLSLVGSST 315
Qy 408 SSALALIIPPLEVTTFFYSEGMSPLTIKDALISILGFGVGVGTYEALYEL 459
Db 316 CCVLGFLPALFHLVLF-KEEMGLQWSSDTAIVVLGVVLAVSGTWSSLSSEI 366
```

RESULT 5

```
US-10-449-902-37011
; Sequence 37011, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37011
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-37011
```

```
Query Match 16.4%; Score 405; DB 6; Length 425;
Best Local Similarity 27.6%; Pred. No. 8e-26;
Matches 120; Conservative 81; Mismatches 164; Indels 70; Gaps 13;

Qy 51 QTLHLKGNIGTGLGLPLAVKAGIWMGIPISILLIIGIVAVHGMGILVKAHFCRL- 109
Db 39 KTFANVFIAVVGSGVLGLPYTFTSRGTGWAGSVLLLAVALTFHCCMLLVACR- 94
Qy 110 --NKSFDYDGTVMYGLSSPCSLRNHAHGRVVDVDFLIYVTLQGFCCVYFVFLADNFK 167
Db 95 YDHPKIASFGD-----LGAAVC-----GPAGRHVVDAMLVLSQASFCVGLIFISNTMA 143
Qy 168 QVIEAANGTTNNCHNNETVILTPMDSR-LYMLSFLPFLVLLVFINRLALSIFSLANI 226
Db 144 HLYPVGDSPPS-----SPLLTAKAIFIWMLPFQGLNSIKITLTLAPLSIFADV 193
Qy 227 TMLVSL-VMYQIVQRIIDPDSHPLVAPW---KTYPLF-----FGTAIFSEFGI 272
Db 194 VDLGAMGVVLGQD-----VSTWLANKPPVPFASAGPTIELYGLGVAVYAFEGI 240
Qy 273 GMVLPLENKKMDPKRPFLLILYLGMIWITLYISLGLCLYQFGANIGSITLNLNPNCLWY 332
Db 241 GMVLPLEAERADKRFKFGTGLALSMAFIAMVYGLFGAMGYLAFAGATRDIIITNLGTGWS 300
Qy 333 QSVKLLYSIGIFTYALQF---YVPAEIIIPFVSRAPEHCELVVDLFRVTVLVCITCL 389
Db 301 VTQVLGCLINLFFTPMVMNPVYEAERLL-----CRKYAWMLRWLLVMVVGML 350
Qy 390 AILPRLDLVLISLVGSSSSALALIIPPLEVTTFFYSEGMSPLTIKDALISILGFGVGFV 449
Db 351 AMLVPNFADFLSLVGSVSVLLGFLVLPAAFLKLVFGAEVGP-GLAGDVAIVVVGTLAV 409
```

```
Qy 273 GMVLPLENKKMDPKRPFLLILYLGMIWITLYISLGLCLYQFGANIGSITLNLNPNCLWY 332
Db 241 GMVLPLEAERADKRFKFGTGLALSMAFIAMVYGLFGAMGYLAFAGATRDIIITNLGTGWS 300
Qy 333 QSVKLLYSIGIFTYALQF---YVPAEIIIPFVSRAPEHCELVVDLFRVTVLVCITCL 389
Db 301 VTQVLGCLINLFFTPMVMNPVYEAERLL-----CRKYAWMLRWLLVMVVGML 350
Qy 390 AILPRLDLVLISLVGSSSSALALIIPPLEVTTFFYSEGMSPLTIKDALISILGFGVGFV 449
Db 351 AMLVPNFADFLSLVGSVSVLLGFLVLPAAFLKLVFGAEVGP-GLAGDVAIVVVGTLAV 409
Qy 450 VGTYEALYELIOPSN 464
Db 410 SGTWTSQAQIFSSD 424
```

RESULT 6

```
US-10-449-902-49196
; Sequence 49196, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49196
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-49196
```

```
Query Match 16.4%; Score 405; DB 6; Length 425;
Best Local Similarity 27.6%; Pred. No. 8e-26;
Matches 120; Conservative 81; Mismatches 164; Indels 70; Gaps 13;

Qy 51 QTLHLKGNIGTGLGLPLAVKAGIWMGIPISILLIIGIVAVHGMGILVKAHFCRL- 109
Db 39 KTFANVFIAVVGSGVLGLPYTFTSRGTGWAGSVLLLAVALTFHCCMLLVACR- 94
Qy 110 --NKSFDYDGTVMYGLSSPCSLRNHAHGRVVDVDFLIYVTLQGFCCVYFVFLADNFK 167
Db 95 YDHPKIASFGD-----LGAAVC-----GPAGRHVVDAMLVLSQASFCVGLIFISNTMA 143
Qy 168 QVIEAANGTTNNCHNNETVILTPMDSR-LYMLSFLPFLVLLVFINRLALSIFSLANI 226
Db 144 HLYPVGDSPPS-----SPLLTAKAIFIWMLPFQGLNSIKITLTLAPLSIFADV 193
Qy 227 TMLVSL-VMYQIVQRIIDPDSHPLVAPW---KTYPLF-----FGTAIFSEFGI 272
Db 194 VDLGAMGVVLGQD-----VSTWLANKPPVPFASAGPTIELYGLGVAVYAFEGI 240
Qy 273 GMVLPLENKKMDPKRPFLLILYLGMIWITLYISLGLCLYQFGANIGSITLNLNPNCLWY 332
Db 241 GMVLPLEAERADKRFKFGTGLALSMAFIAMVYGLFGAMGYLAFAGATRDIIITNLGTGWS 300
Qy 333 QSVKLLYSIGIFTYALQF---YVPAEIIIPFVSRAPEHCELVVDLFRVTVLVCITCL 389
Db 301 VTQVLGCLINLFFTPMVMNPVYEAERLL-----CRKYAWMLRWLLVMVVGML 350
Qy 390 AILPRLDLVLISLVGSSSSALALIIPPLEVTTFFYSEGMSPLTIKDALISILGFGVGFV 449
Db 351 AMLVPNFADFLSLVGSVSVLLGFLVLPAAFLKLVFGAEVGP-GLAGDVAIVVVGTLAV 409
```

```
QY 450 VGTAEALYELQPSN 464
Db 410 SGTWTSQAIFSSSD 424

RESULT 7
US-10-449-902-32842
; Sequence 32842, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; *PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32842
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-32842

Query Match 16.4%; Score 404.5; DB 6; Length 422;
Best Local Similarity 26.9%; Pred. No. 8.7e-26;
Matches 126; Conservative 79; Mismatches 186; Indels 77; Gaps 13;

QY 28 GLNNLSSPGSYQRF-----GQSNSTTWFTQTLIHLKGNIGTGLGLPLAVKNA 75
Db 2 GLGNEASSSSRLDPAFLPHHGGDGAGKLSQPKTFANFVAVVGAGVGLGPLYTFST 61
QY 76 GIVMGPSILLIGIVAVHCGILVCAHHCRRLNK--SFVDYGDVTMYGLESSPCSWLR 133
Db 62 GWAAGSILLISVAALTYFCMMLLVACRRRLADEHPKIASFGDLGDAVFRG----P----- 112
QY 134 NHAHWRGVVDFLIIVTQLGFCVYFVLADNFKQVIEAANGTNNCHNETVILPTMD 193
Db 113 -----GRLAVDTMVLQSASFVGYLIFISNTWAHLYPVFPAPSSN-----ALLSP--- 157
QY 194 SRLYMLSFLLPVLVFIIRNLRAISFSLLANITMLVSLVMYIQFIVQRIIPDPSPHPLVA 253
Db 158 KALFIWAMLPFQLGLNSIKTLTLLAPLSIFADVVDLGAMGVILG-----EDVS 205
QY 254 PW--KTYPLF-----FCTAIFSEGIGMVLPLENKMKDPKFFLILVGMVIT 300
Db 206 VWIAKPPVFAFGLSAILYIGVSVYAFEGIGMVLPLEAANKKFGTTGLSMGFIA 265
QY 301 ILVYISLCLGYLQFAGNIQGSITLNLPCWLYOSVLLYSIGIFFTYALQF---YVPAEI 357
Db 266 VMYGLFGAMGYIAGDATRDIITNLGTWLSAAVQLGCLINLFFTPVMVHPYVEAER 325
QY 358 IIPFFVSRAPHEHCELVDLIVRTVLVCLTCILAILPRDLVLSVGSVSSSALALIIPP 417
Db 326 LL-----HGKRYC-----WMLRWLLVAVGLSAMYPNFTDFLALVGSVSVCLLGFVLPA 375
QY 418 LLEVTTYSEGMSPLTIFKDALISILGFGVGVVGTVEALYELQPSNA 465
Db 376 SFHLKVFGE-MSWSGVLSVLLVLLGLSLAVFGTYSLLQIFHSSSA 422

RESULT 8
US-10-449-902-53550
; Sequence 53550, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; *PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53550
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-53550

Query Match 16.4%; Score 404.5; DB 6; Length 422;
Best Local Similarity 26.9%; Pred. No. 8.7e-26;
Matches 126; Conservative 79; Mismatches 186; Indels 77; Gaps 13;

QY 28 GLNNLSSPGSYQRF-----GQSNSTTWFTQTLIHLKGNIGTGLGLPLAVKNA 75
Db 2 GLGNEASSSSRLDPAFLPHHGGDGAGKLSQPKTFANFVAVVGAGVGLGPLYTFST 61
QY 76 GIVMGPSILLIGIVAVHCGILVCAHHCRRLNK--SFVDYGDVTMYGLESSPCSWLR 133
Db 62 GWAAGSILLISVAALTYFCMMLLVACRRRLADEHPKIASFGDLGDAVFRG----P----- 112
QY 134 NHAHWRGVVDFLIIVTQLGFCVYFVLADNFKQVIEAANGTNNCHNETVILPTMD 193
Db 113 -----GRLAVDTMVLQSASFVGYLIFISNTWAHLYPVFPAPSSN-----ALLSP--- 157
QY 194 SRLYMLSFLLPVLVFIIRNLRAISFSLLANITMLVSLVMYIQFIVQRIIPDPSPHPLVA 253
Db 158 KALFIWAMLPFQLGLNSIKTLTLLAPLSIFADVVDLGAMGVILG-----EDVS 205
QY 254 PW--KTYPLF-----FCTAIFSEGIGMVLPLENKMKDPKFFLILVGMVIT 300
Db 206 VWIAKPPVFAFGLSAILYIGVSVYAFEGIGMVLPLEAANKKFGTTGLSMGFIA 265
QY 301 ILVYISLCLGYLQFAGNIQGSITLNLPCWLYOSVLLYSIGIFFTYALQF---YVPAEI 357
Db 266 VMYGLFGAMGYIAGDATRDIITNLGTWLSAAVQLGCLINLFFTPVMVHPYVEAER 325
QY 358 IIPFFVSRAPHEHCELVDLIVRTVLVCLTCILAILPRDLVLSVGSVSSSALALIIPP 417
Db 326 LL-----HGKRYC-----WMLRWLLVAVGLSAMYPNFTDFLALVGSVSVCLLGFVLPA 375
QY 418 LLEVTTYSEGMSPLTIFKDALISILGFGVGVVGTVEALYELQPSNA 465
Db 376 SFHLKVFGE-MSWSGVLSVLLVLLGLSLAVFGTYSLLQIFHSSSA 422

RESULT 9
US-10-449-902-53486
; Sequence 53486, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; *PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53550
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-53550
```

;; PRIOR FILING DATE: 2002-12-11
;; NUMBER OF SEQ ID NOS: 56791
;; SOFTWARE: PatentIn ver. 2.1
;; SEQ ID NO 53486
;; LENGTH: 422
;; TYPE: PRT
;; ORGANISM: Oryza sativa
US-10-449-902-53486

Query Match 16.3%; Score 402.5; DB 6; Length 422;
Best Local Similarity 26.7%; Pred. No. 1.3e-25;
Matches 124; Conservative 79; Mismatches 191; Indels 71; Gaps 12;
Qy 28 GLNNLSPPGSYQRF-----GOSNTTWFTQTLIHLKGNIGTGLGLPLAVKNA 75
Db 2 GLGNEASSSSKRLDPAPLLPHGGDAGKLSQPKTFANFVIAVVGAGVLGLPVTFSRT 61
Qy 76 GIVMGPSILLIIGIVAVHGMGILVKAHFRCRLNK--SFVDYGDVTVMYGLESSPCSWLR 133
Db 62 GWAAGSILLLSVAALTFFYCMMLLVACRRRLADEHPKIASFGDLGDAVFRG----P----- 112
Qy 134 NHAHWGRVVDFFLIVTQGFCCVYFVFLADNFKQVIEAANGTTNNCHNETVILTPMD 193
Db 113 -----GRLAVDTMLVLSQASECVGYLIFISIMAHLYPVFPAPSN-----ALLSP--- 157
Qy 194 SRLYMLSPFLPVLVLFIRNLRALSIFSLLANITMLVSLVMYQFIVQRIIDPPSHLPLVA 253
Db 158 KALFIWMLPQLGLNSIKTITLLAPLSIFADVVDLGAMGVVG-----EDVS 205
Qy 254 PW--KTVPFP-----FGTAIFSPGIGMVLPLENKKDKPRKFPPLILYLGWIVT 300
Db 206 VMLAKPPVPVAFGGLSAILYIGVSVTAFEGIGMVLPLEAEAAANKKRGTTGLSLMGFIA 265
Qy 301 ILYISLGLVLOFGANIQSITLNPNCWLYQSVKLYSIGIFETVALQFYVPAEIIIP 360
Db 266 VMYGLFGAMGYIAFGDATRDIITNLGTGWSAAVQLGCLNLFPTMPVMHPYVEVAKR 325
Qy 361 FFVSRAPCHCLVVDLFFVTVLCLTCILAILPRDLVLSVSSSALALAIIPPLLE 420
Db 326 LL--HGKRYC-----WMLRWLLVLAVGLSAMYVNFDTDFLALVGSSVCVLLGFLVPASFH 378
Qy 421 VTFYSGMSPLTIKDALISILFGVGVVGTVEALYELIQPSNA 465
Db 379 LKVFGEA-MSWSGVLSVDLLVLLGLSLAVFGTYTSLLIQIFHSSA 422

RESULT 10
US-10-449-902-29254
; Sequence 29254, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agricultural Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205V1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 29254
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-29254

Query Match 16.2%; Score 400; DB 6; Length 395;
Best Local Similarity 27.6%; Pred. No. 1.9e-25;

Matches 116; Conservative 75; Mismatches 158; Indels 72; Gaps 11;
Qy 51 QTLIHLKGNIGTGLGLPLAVKNAIGVMGPISLIIGIVAVHGMGILVKAHFRCRL- 109
Db 39 KTFANFVIAVVGSGVLGLPFTFSRTGWAGSVLLAVAAALTFHGMMLVACR---RRLA 94
Qy 110 --NKSFDYDGTVMYGLESSPCSWLRNHAHWGRVVDFFLIVTQGFCCVTVFVLADNPK 167
Db 95 YDHPKIASFGD-----LGAAYC-----GPAGRHVVDAMLVLSQASFVGLYIFISNTMA 143
Qy 168 QVIEAANGTTNNCHNETVILTPMDSR-LYMLSPFLPVLVLFIRNLRALSIFSLLANI 226
Db 144 HLYPVGSSPB-----SPLUTAKAIFWMLPQLGLNSIKTITLLAPLSIFAPG 193
Qy 227 TMLVSLVMYQFIVQRIIDPPSHLPLVAPWKTYPLFFGTAFISFEGIGMVLPLENKKDKPR 286
Db 194 TEIL-----YGL- -GVAVYAFEGIGMVLPLEAEAAADR 224
Qy 347 YALQF---YVPAEIIIPFFVSRAPCHCLVVDLFFVTVLCLTCILAILPRDLVLSLV 403
Db 285 MPVMNPNPYEVAERLL-----CKRYAWMLRWLLVMVGLMAMLVNPFADFLSLV 334
Qy 404 GSVSSSALALAIIPPLLEVTTFYSGMSPLTIKDALISILFGVGVVGTVEALYELIQPS 463
Db 335 GSSVCVLLGFLVPAAPHLKVFGEVGP-GLAGDVAVIVVGTALAVSGTWTSLAQIFSSS 393
Qy 464 N 464
Db 394 D 394

RESULT 11
US-10-953-349-370
; Sequence 370, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 370
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-370

Query Match 15.5%; Score 382; DB 6; Length 355;
Best Local Similarity 32.8%; Pred. No. 5.1e-24;
Matches 122; Conservative 51; Mismatches 131; Indels 68; Gaps 12;
Qy 112 SFVDYDGTVMYGLESSPCSWLRNHAHWGRVVDFFLIVTQGFCCVTVFVLADNPKQVIE 171
Db 23 SFGDLGFAV-----CGSL-----GRIVVDLFIILSQAGFCVGYLIFI----- 59
Qy 172 AANGTT--NNCHNETVILTP-----MDSRLYMLSPFLPVLVLFIRNL 214
Db 60 ---GTTLANSDPE-----SPTSLRHQTRLGSEFLGVSSKSLYINGCFFQLGLNSIKTL 112
Qy 215 RALSIFSLLANI-----TMLVSLVMYQFIVQRIIDPPSHLPLVAPWKTYPLF---FGTAIF 267
Db 113 THLAPLSIFADIVDLGAMAVVIVEDSMILKQRPD-----VVAFGGMSFLYGMGVAVY 166
Qy 268 SFEGIGMVLPLENKKDKPRKFPPLILYLGWIVTILYISLGLVLOFGANIQSITLNL 327
Db 167 SFEGVGWVLPLESEMKKDKFKGLALGMGFIISLYIAFGILGYLAFGEDTMDITIANLG 226

LENGTH: 531
TYPE: Prt
ORGANISM: Triticum aestivum
FEATURE:
NAME/KEY: misc_feature
LOCATION:
OTHER INFORMATION: Pfam Name: Aa_trans; Pfam Description: Transmembrane amino
FEATURE:
NAME/KEY: misc_feature
LOCATION:
OTHER INFORMATION: GI Number: 50904567; NR Description: putative amino acid
OTHER INFORMATION: transport protein [Oryza sativa (japonica cultivar-group)]
OTHER INFORMATION: >gi|41053220|dbj|BAD08181.1| putative amino acid transport protei
FEATURE:
NAME/KEY: misc_feature
LOCATION:
OTHER INFORMATION: GI Number: 50915446; NR Description: amino acid
OTHER INFORMATION: transporter-like [Oryza sativa (japonica cultivar-group)]
OTHER INFORMATION: >gi|47497045|dbj|BAD13097.1| amino acid transporter-like [Oryza
OTHER INFORMATION: sativa (japonica cultivar-group)] >gi|47497767|dbj|BAD19867.1|
FEATURE:
NAME/KEY: misc_feature
LOCATION:
OTHER INFORMATION: GI Number: 30687791; NR Description: amino acid transporter
FEATURE:
NAME/KEY: misc_feature
LOCATION:
OTHER INFORMATION: GI Number: 42563875; NR Description: amino acid transporter
FEATURE:
NAME/KEY: misc_feature
LOCATION:
OTHER INFORMATION: GI Number: 42563873; NR Description: amino acid transporter
US-11-174-307B-5170

Query Match 12.9%; Score 318; DB 7; Length 531;
Best Local Similarity 25.8%; Pred. No. 1.6e-18;
Matches 115; Conservative 73; Mismatches 175; Indels 82; Gaps 18;
Qy 42 GQSTMTFOTLIHLKNGITGGLPLAVKNAGVMPISLLIIGIVAVHC--MGILV 99
Db 140 GQGSVT--QTVFNGINVLAVGLLSAPFTIHEAGWA---GLAVLSVFAIICCYTGVL 193
Qy 100 KCAHFHRRLNKSFVDYGDVTMYGLESPCSWLRNHAHWGR--RVDPFLIVTQGFCCV 157
Db 194 K---HCFESKDGATYDI-----GEAFGRIGRLIISILVTELYSYCV 235
Qy 158 YFVFL-ADNFQVTEAANGTNNCHNETVILPTMDSRRLYMLSPFLPLVLL--VFIRNL 214
Db 236 EFIILEGDNMTSIFPDVNINLFGIH-----VDSK-HFFGVLTALVVLPTVWLRLD 284
Qy 215 RALSIFS---LLANITMLVSLVMI-----YQFVQRI DPDPHPLVAPWKTYPLPFG 263
Db 285 RVLSYLSAGGVATLWFLSVALVGTTEGIGFHO-----TGSAVKWSGMPFAIG 333
Qy 264 TAIFSEFEGIMVPLENKMDPRKPLIYLGWIVITLYISLGLGYLQFGANIQGSIT 323
Db 334 IYGCYSGHVPFNIYQMSDRTPFKALFICFAVCTAIGSFAIIGYLMFGDKTLSQIT 393
Qy 324 LNLPCMLYQSVKLLYS-IGIFFTVALQFYVPAEIIIPFFVSRAPHC-----ELVVD 375
Db 394 LNLKESFASKVALWTVINPFTKAL-----LLNP--LARSLEELRPEGFLNETIVS 444
Qy 376 LFRVTVLCLTILAILPRDLVLISGVSSSALALIIPPLEVITFFYSEGMSPLTIF 435
Db 445 IILRTSLVASTVIAFLPFFGLVNLIGLSLILVAVIMPALCLFKIAQNKAATRPQVIA 504
Qy 436 KDALISILGFVGVVGTVEALVELI 460
Db 505 SVAII-VGVVISAAIGTYSVASII 528

RESULT 15

US-11-056-355B-105677

Sequence 105677, Application US/11056355B
Publication No. US20060150283A1
GENERAL INFORMATION:
APPLICANT: Brover, Vvacheslav
TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
FILE OF INVENTION: Polypeptides Encoded Thereby
FILE REFERENCE: 2750-1590PUS2
CURRENT APPLICATION NUMBER: US/11/056.355B
CURRENT FILING DATE: 2005-02-14
PRIOR APPLICATION NUMBER: 60/544,190
PRIOR FILING DATE: 2004-02-13
NUMBER OF SEQ ID NOS: 119966
SEQ ID NO 105677
LENGTH: 468
TYPE: prt
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: peptide
LOCATION: (1)..(468)
OTHER INFORMATION: Ceres Seq. ID no. 13617748
US-11-056-355B-105677

Query Match 12.2%; Score 302; DB 7; Length 468;
Best Local Similarity 25.3%; Pred. No. 2.9e-17;
Matches 123; Conservative 87; Mismatches 188; Indels 88; Gaps 20;
Qy 2 STQRLNEDYHYSST--DVSPEPSSEGLNLSPPGSYQRFQGSNSTTWFQTLIHLK 59
Db 37 SSKPLLSQPPDKETILPVNPQSLKSLVTDLP-----EPNLCSFSQSVLNGTIV 89
Qy 60 NIGTGLGLPLAVKNAGVMPISLLIIGIVAVHCMLVKCAHFHRRLNK-----FV 114
Db 90 LCGLGLITMPVAIKESGWLGLPI-LLPFGVITCY-TGVLMK-----RCLESSPGIQTY 141
Qy 115 DYGTVMYGLESPCSWLRNHAHWGRVDFLIVTQGFCCVYFVFLADN----FKOV- 169
Db 142 DIGQAA-FGIT-----GRFIISILLVELYAAACVEYIIMSDNLSGLFPNVS 187
Qy 170 IEAANGTNNCHNETVILPTMDSRRLYMLSPFLPLVLLVFIIRNLRLAISIFS---LLANI 226
Db 188 LSIASGIS-----LDSQIFAILTTLVLP---TWLKDLSLLSYLVSGVGLASI 234
Qy 227 TMLVSLVMIYQ-----FIVQRI DPDPHPLVAPWKTYPLPFGTAIFSFEIGIMVLP 280
Db 235 LLGICLEWGVAVDVGIGFHTATGRVFDLSNLPV-----IGIFGFGYSGHSPFPNIY 285
Qy 281 KMKDPRKPLIYLGWIVITLYISLGLGYLQFGANIQGSITLNLPCMLYQSVKLLYS 340
Db 286 SMKDPSPRPPLVLVIFCFCTVLYIAVAVCGYTMFEAVESQFTLNMKPHFFPSKVAVMTA 345
Qy 341 IGIFFT-YAL--QFYVPAEIIIPFFVSRAPHCVELVVDLFRVTVLCLTILAILPRL 396
Db 346 VITPMTKALTITPVMKSLLELIPTAKMSRG-----VSILFRMLVTLSTLVVALSVFF 400
Qy 397 DLVLSLVGSSSSALALIIPPLEVITFF---YSEGMSPLTIFKDALISILGFVGVFVQTY 453
Db 401 AIVAAALIGSFLAMLVALIFPCLCYLSILKGLSNTQIGLCIF---IIVFGVWSGCCGT 456
Qy 454 EALVEL 459
Db 457 SAISRL 462

Search completed: August 24, 2006, 01:30:25

Job time : 34 secs

THIS PAGE BLANK (USPTO)